Jürgen Audretsch

Entangled Systems

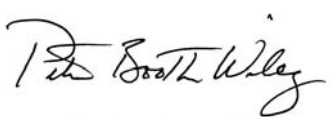
1807–2007 Knowledge for Generations

Each generation has its unique needs and aspirations. When Charles Wiley first opened his small printing shop in lower Manhattan in 1807, it was a generation of boundless potential searching for an identity. And we were there, helping to define a new American literary tradition. Over half a century later, in the midst of the Second Industrial Revolution, it was a generation focused on building the future. Once again, we were there, supplying the critical scientific, technical, and engineering knowledge that helped frame the world. Throughout the 20th Century, and into the new millennium, nations began to reach out beyond their own borders and a new international community was born. Wiley was there, expanding its operations around the world to enable a global exchange of ideas, opinions, and know-how.

For 200 years, Wiley has been an integral part of each generation's journey, enabling the flow of information and understanding necessary to meet their needs and fulfill their aspirations. Today, bold new technologies are changing the way we live and learn. Wiley will be there, providing you the must-have knowledge you need to imagine new worlds, new possibilities, and new opportunities.

Generations come and go, but you can always count on Wiley to provide you the knowledge you need, when and where you need it!

William J. Pesce
President and Chief Executive Officer



Peter Booth Wiley
Chairman of the Board

Jürgen Audretsch

Entangled Systems

New Directions in Quantum Physics

WILEY-VCH Verlag GmbH & Co. KGaA

The Author

Jürgen Audretsch
Fachbereich Physik
Universität Konstanz
E-mail: juergen.audretsch@uni-konstanz.de

Cover illustration

Itten, Johannes: Die Begegnung, 1916
© VG Bild-Kunst, Bonn 2006

Original Title:
Verschränkte Systeme – Die Quantenphysik auf neuen Wegen © 2005 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. The German edition has been revised and extended before translation.

Translated from the German by Prof. William D. Brewer, Freie Universität Berlin, Germany.

All books published by Wiley-VCH are carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained in these books, including this book, to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Library of Congress Card No.:
applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

© 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Typesetting Uwe Krieg, Berlin
Printing Strauss GmbH, Mörlenbach
Binding Litges & Dopf GmbH, Heppenheim
Wiley Bicentennial Logo Richard J. Pacifico

Printed in the Federal Republic of Germany
Printed on acid-free paper

ISBN-13: 978-3-527-40684-5
ISBN-10: 3-527-40684-0

Contents

Preface to the English Edition **XIII**

Preface to the German Edition **XV**

1 The Mathematical Framework **1**
1.1 Hilbert Vector Space . . . 2
1.1.1 The Scalar Product and the Dirac Notation . . . 2
1.1.2 Linear Operators on the Hilbert Space . . . 3
1.1.3 Normal Operators and Spectral Decompositions . . . 6
1.1.4 Hermitian Operators . . . 9
1.1.5 Unitary Operators . . . 11
1.1.6 Positive Operators and Projection Operators . . . 11
1.2 Liouville Operator Space . . . 13
1.2.1 Scalar Product . . . 13
1.2.2 Superoperators . . . 14
1.3 The Elements of Probability Theory . . . 15
1.3.1 The Probability of Random Events . . . 15
1.3.2 Conditional Probability and Bayes' Theorem . . . 17
1.3.3 Random Quantities . . . 19
1.4 Complementary Topics and Further Reading . . . 19
1.5 Problems for Chapter 1 . . . 20

2 Basic Concepts of Quantum Theory **23**
2.1 First Version of the Postulates (pure states of isolated quantum systems) . . . 23
2.1.1 Introduction: the Scenario of Quantum Mechanics . . . 23
2.1.2 Quantum States . . . 29
2.1.3 Postulates for Pure States of Isolated Quantum Systems . . . 32
2.1.4 Comments on the Postulates . . . 35
2.2 Outlook . . . 37
2.3 Manipulation of the Evolution of the States by Projective Measurements . . . 38
2.3.1 The Quantum Zeno Effect . . . 38
2.3.2 Driving a State Vector by a Sequence of Projection Measurements . . . 39
2.4 The Structure of Physical Theories* . . . 40

The chapters marked with an asterisk * can be skipped over in a first reading.

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

2.4.1 Structural Elements of a Physical Theory* . . . 41
2.4.2 Developed Reality* . . . 42
2.5 Interpretations of Quantum Theory and Physical Reality* . . . 43
2.5.1 The Minimal Interpretation* . . . 43
2.5.2 The Standard Interpretation* . . . 44
2.6 Complementary Topics and Further Reading . . . 46

3 The Simplest Quantum Systems: Qubits 49
3.1 Pauli Operators . . . 50
3.2 Visualisation of Qubits on the Bloch Sphere . . . 52
3.3 Visualisation of the Measurement Dynamics and the Unitary Dynamics . . . 55
3.4 Quantum Gates for Single Qubit Systems . . . 59
3.5 Spin-$\frac{1}{2}$ . . . 62
3.6 Photon Polarisations . . . 62
3.7 Single Photons in a Beam Splitter and in an Interferometer . . . 63
3.7.1 Beam Splitters . . . 64
3.7.2 Interferometer . . . 66
3.8 Locating a Bomb Without Exploding It by Using a Null Measurement* . . . 68
3.9 Complementary Topics and Further Reading . . . 71
3.10 Problems for Chapter 3 . . . 71

4 Mixed States and the Density Operator 73
4.1 Density Operators for a Given Ensemble (Statistical Mixture) . . . 73
4.1.1 Pure States . . . 73
4.1.2 The Physics of Statistical Mixtures (Blends) . . . 75
4.1.3 Definition and Properties of the Generalised Density Operator . . . 78
4.1.4 Incoherent Superpositions of Pure States . . . 80
4.2 The Generalised Quantum State . . . 82
4.3 Different Ensemble Decompositions of a Density Operator and the Ignorance Interpretation . . . 82
4.4 Density Operators of Qubits . . . 85
4.5 Complementary Topics and Further Reading . . . 86
4.6 Problems for Chapter 4 . . . 86

5 Shannon's Entropy and Classical Information 89
5.1 Definition and Properties . . . 89
5.2 Shannon's Theorem . . . 93
5.2.1 Typical Sequences . . . 93
5.2.2 Classical Data Compression . . . 95
5.3 Classical Information . . . 96
5.4 Classical Relative Entropy . . . 97
5.5 Mutual Information as a Measure of the Correlation between Two Messages . . . 97
5.5.1 Mutual Information . . . 98
5.5.2 Conditional Entropy . . . 99
5.6 Complementary Topics and Further Reading . . . 101
5.7 Problems for Chapter 5 . . . 101

6 The von Neumann Entropy and Quantum Information **103**
6.1 The Quantum Channel and Quantum Entropy . . . 103
6.2 Qubits as the Unit of Quantum Information . . . 106
6.3 Properties . . . 108
6.4 The Interfaces of Preparation and Measurement . . . 110
6.4.1 The Entropy of Projective Measurements . . . 110
6.4.2 The Entropy of Preparation . . . 111
6.5 Quantum Information . . . 112
6.6 Complementary Topics and Further Reading . . . 112
6.7 Problems for Chapter 6 . . . 113

7 Composite Systems **115**
7.1 Subsystems . . . 115
7.2 The Product Hilbert Space . . . 116
7.2.1 Vectors . . . 116
7.2.2 Operators . . . 118
7.3 The Fundamentals of the Physics of Composite Quantum Systems . . . 120
7.3.1 Postulates for Composite Systems and Outlook . . . 120
7.3.2 The State of a Subsystem, the Reduced Density Operator, and General Mixtures . . . 122
7.4 Manipulations on a Subsystem . . . 124
7.4.1 Relative States and Local Unitary Transformations . . . 124
7.4.2 Selective Local Measurements . . . 125
7.4.3 A Non-Selective Local Measurement . . . 127
7.5 Separate Manipulations on both Subsystems . . . 128
7.5.1 Pairs of Selective Measurements . . . 128
7.5.2 Non-Local Effects: "Spooky Action at a Distance"? . . . 130
7.6 The Unitary Dynamics of Composite Systems . . . 131
7.7 A First Application of Entanglement: a Conjuring Trick . . . 132
7.7.1 The Conjuring Trick . . . 132
7.7.2 Classical Correlations can give no Explanation . . . 133
7.7.3 The Trick . . . 134
7.8 Quantum Gates for Multiple Qubit Systems . . . 135
7.8.1 Entanglement via a CNOT Gate . . . 135
7.8.2 Toffoli, SWAP, and Deutsch Gates . . . 137
7.9 Systems of Identical Particles* . . . 138
7.10 Complementary Topics and Further Reading . . . 141
7.11 Problems for Chapter 7 . . . 142

8 Entanglement **143**
8.1 Correlations and Entanglement . . . 143
8.1.1 Classically-Correlated Quantum States and LOCC . . . 143
8.1.2 Separability and Entanglement . . . 145
8.1.3 The Separability Problem . . . 147
8.2 Outlook . . . 148

8.3 Entangled Pure States . . . 149
8.3.1 The Schmidt Decomposition . . . 149
8.3.2 The Schmidt Number and Entanglement . . . 151
8.3.3 The Entropy of the Subsystems as a Measure of Entanglement . . . 152
8.3.4 Subsystems in Pure States are not Entangled* . . . 153
8.4 The PPT Criterion for the Entanglement of Mixtures * . . . 155
8.5 The Production of Entangled States . . . 157
8.6 The No-Cloning Theorem Prevents Transfer of Information Faster than the Velocity of Light . . . 159
8.7 Marking States by Entanglement* . . . 161
8.7.1 Which-Way Marking* . . . 161
8.7.2 Quantum Erasure* . . . 164
8.7.3 Delayed Choice of the Marker Observables* . . . 165
8.8 Complementary Topics and Further Reading . . . 166
8.9 Problems for Chapter 8 . . . 167

9 Correlations and Non-Local Measurements 169
9.1 Entropies and the Correlations of Composite Quantum Systems . . . 169
9.1.1 Mutual Information as a Measure of Correlations . . . 169
9.1.2 The Triangle Inequality . . . 170
9.1.3 Entangled *vs.* Classically-Correlated Quantum Systems . . . 171
9.2 Non-Local Measurements . . . 174
9.2.1 The Bell Basis . . . 174
9.2.2 Local and Non-Local Measurements . . . 175
9.2.3 Non-local Measurements by Means of Local Measurements on an Ancillary System . . . 177
9.2.4 Non-locally Stored Information and Bell Measurements . . . 179
9.3 Complementary Topics and Further Reading . . . 181
9.4 Problems for Chapter 9 . . . 181

10 There is no (Local-Realistic) Alternative to Quantum Theory 183
10.1 EPR Experiments and Their Quantum-mechanical Explanation . . . 183
10.2 Correlated Gloves . . . 186
10.3 Local Realism . . . 187
10.4 Hidden Variables, Bell Inequalities and Contradictions of Experiments . . . 188
10.5 Separable Mixtures Obey the Bell Inequality . . . 191
10.6 Entanglement Witnesses* . . . 192
10.7 3-Particle Entanglement and Quantum Locality . . . 194
10.7.1 The GHZ State . . . 194
10.7.2 Local Realism and Quantum Theory at Odds . . . 194
10.8 Complementary Topics and Further Reading . . . 196
10.9 Problems for Chapter 10 . . . 197

11 Working with Entanglement **199**
11.1 Quantum Cryptography . . . 199
11.1.1 The Vernam Coding . . . 199
11.1.2 The B92 Protocol . . . 200
11.1.3 EPR Protocols . . . 202
11.1.4 The Scheme of Quantum Cryptography . . . 203
11.2 One Qubit Transmits Two Bits (Dense Coding) . . . 204
11.3 Quantum Teleportation . . . 204
11.4 Entanglement Swapping . . . 207
11.5 Spooky Action into the Past?* . . . 208
11.6 Entanglement Distillation . . . 209
11.7 A Measure of Entanglement for Mixtures: Entanglement of Formation and Concurrence* . . . 212
11.8 Complementary Topics and Further Reading . . . 216
11.9 Problems for Chapter 11 . . . 217

12 The Quantum Computer **219**
12.1 Registers and Networks . . . 219
12.2 Functional Computation . . . 221
12.3 Quantum Parallelism . . . 224
12.4 Two Simple Quantum Algorithms . . . 225
12.4.1 The Deutsch Problem . . . 225
12.4.2 The Deutsch-Jozsa Problem . . . 226
12.5 Grover's Search Algorithm . . . 228
12.6 Shor's Factorisation Algorithm . . . 230
12.6.1 Reduction of Factorisation to the Search for a Period . . . 231
12.6.2 The Quantum Algorithm for Determining the Period . . . 234
12.7 Quantum Error Correction Using Non-local Measurements . . . 239
12.7.1 Bit Flip Errors . . . 239
12.7.2 Phase Flip Errors . . . 240
12.8 The Components of the Quantum Computer* . . . 241
12.9 Complementary Topics and Further Reading . . . 243
12.10 Problems for Chapter 12 . . . 244

13 Generalised Measurements, POVM **247**
13.1 The Function of a Generalised Dynamics of Open Quantum Systems . . . 247
13.1.1 Problems . . . 247
13.1.2 A Simple Example . . . 248
13.2 The Non-optimal Stern-Gerlach Experiment and Generalised Measurements . . . 251
13.2.1 The Experimental Setup . . . 251
13.2.2 An Example of a Generalised Measurement . . . 253
13.2.3 Unsharp Measurements . . . 255
13.3 Generalised Measurements . . . 256
13.3.1 What is a Quantum Measurement? . . . 256
13.3.2 Generalised Measurement Postulates . . . 257

13.3.3 The Polar Decomposition of a Linear Operator . . . 258
13.3.4 Minimal Measurements and POVM . . . 260
13.3.5 Implementation of a Generalised Measurement by Unitary Transformation and Projection . . . 261
13.3.6 Entanglement Distillation by means of Generalised Measurements* . . . 262
13.4 POVM Measurements . . . 263
13.4.1 Measurement Probabilities and Positive Operators . . . 263
13.4.2 A Composite Measurement as an Example of a POVM Measurement 264
13.4.3 Can One Distinguish Between Two States with Certainty by a Single POVM Measurement? . . . 265
13.4.4 The Advantage of a POVM Measurement for Determining States . . . 267
13.4.5 An Informationally-Complete POVM* . . . 268
13.4.6 Estimating the State Before the Measurement . . . 269
13.5 Complementary Topics and Further Reading . . . 270
13.6 Problems for Chapter 13 . . . 270

14 The General Evolution of an Open Quantum System and Special Quantum Channels 273
14.1 Quantum Operations and their Operator-Sum Decompositions . . . 273
14.1.1 Quantum Operations . . . 273
14.1.2 The Operator-Sum Decomposition of Quantum Operations . . . 275
14.1.3 Simple Quantum Operations . . . 276
14.1.4 The Ambiguity of the Operator-Sum Decomposition . . . 277
14.2 The Master Equation . . . 277
14.3 Completely General Selective Measurements and POVM . . . 279
14.4 Quantum Channels . . . 281
14.4.1 The Depolarising Channel . . . 281
14.4.2 Quantum Jumps and Amplitude Damping Channels . . . 282
14.4.3 An Entanglement-Breaking Channel * . . . 283
14.5 The Scenario and the Rules of Quantum Theory Revisited . . . 284
14.6 Complementary Topics and Further Reading . . . 288
14.7 Problems for Chapter 14 . . . 288

15 Decoherence and Approaches to the Description of the Quantum Measurement Process 289
15.1 Channels which Produce Decoherence . . . 289
15.1.1 The Phase Damping Channel . . . 289
15.1.2 Scattering and Decoherence . . . 291
15.1.3 The Phase Flip Channel . . . 292
15.2 Environment-Induced Decoherence . . . 293
15.2.1 The Formation of the Classical World . . . 293
15.2.2 Schrödinger's Cat . . . 296
15.3 The Quantum Measurement Process* . . . 297
15.3.1 The Research Programme* . . . 297
15.3.2 Pre-Measurement* . . . 298

15.3.3 Entanglement with the Environment Fixes the Observable* 299
15.3.4 Entanglement with Many Degrees of Freedom of the Environment* . 300
15.4 Has the Problem of Measurements been Solved?* 303
15.5 The Many-Worlds Interpretation* 303
15.6 Complementary Topics and Further Reading 304
15.7 Problems for Chapter 15 306

16 Two Implementations of Quantum Operations* **307**
16.1 The Operator-Sum Decomposition* 307
16.2 The Unitary Implementation of Quantum Operations* 310
16.3 Implementation of a Completely General Selective Measurement by Unitary Transformation and Projection* 311
16.4 Complementary Topics and Further Reading* 313
16.5 Problems for Chapter 16 313

References **315**
Reference categories 315
Bibliography 316

Subject Index **331**

Preface to the English Edition

For the English edition, the German text has been revised and extended with a number of new sections. I have taken the suggestions of students into account, who have used the text as an accompaniment to lecture courses and seminars. I wish to express my sincere thanks to them and to all of the others who have discussed the book with me. I am grateful for indications of typographical errors, which have been corrected. Furthermore, I once again wish to thank Michael Nock and Stefan Bretzel for their diligent help with the manuscript.

The famous aphorism "If I have seen further it is by standing on ye sholders of Giants"[1] is due to Sir Isaac Newton. This statement holds for anyone who works in the field of quantum theory today. We are only the dwarves, who owe so much to the giants. However, the question remains as to how the dwarves get up onto the shoulders of the giants as quickly and easily as possible. This book is intended to be of help precisely in this process.

Jürgen Audretsch

Konstanz, November 2006

[1] I. Newton in his letter to R. Hooke of February 5, 1675. Further details can be found in [Mer 65].

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

Preface to the German Edition

This book is a

textbook in theoretical physics.

It is the result of lecture courses and seminars which I have given in recent years at the University of Konstanz, Germany, on the subjects of

quantum information theory and the
foundations of quantum mechanics.

Non-relativistic quantum mechanics has experienced a phase of turbulent development in the past few years. Quantum computing, quantum teleportation, quantum cryptography, and quantum information are typical buzzwords which reflect these developments, and they can be found beyond the professional circle of physics in popular-science articles and in Sunday supplements. The concept of *entanglement* is the central theoretical idea accompanying these "new directions" in quantum mechanics, which are increasingly making their way into physics curricula in schools. The theoretical fundamentals of these new developments are the subject of this book.

For whom is the book intended? The readers targeted are in the main students, but in addition, all those who are interested in quantum mechanics or perhaps simply fascinated by the subject. This includes not only physicists and students of physics, but also students of computer science, chemistry and other natural sciences, as well as engineers and teachers. The book presumes a basic acquaintance with quantum mechanics: it does not start from zero.

To be sure, all the basics of mathematics and physics which are required for understanding the later chapters are summarised in the initial Chapters 1 and 2; they are treated there from a point of view which may be new to many readers. These chapters serve among other things as a preparation for the fact that in quantum mechanics, the concepts of a *state* and the *evolution of a state*, including measurements, are used in a different sense from that of classical physics. The generalisations in the later chapters build on these fundamentals. The second chapter also contains a toolbox for the philosophy of science, with which we can discuss the question of the nature of the reality on which quantum mechanics is based.

Finally, the demands made upon the reader increase from chapter to chapter. The earlier chapters form a basis for the later ones; the exercises at the end of each chapter can serve as a guide for the reader as to whether he or she has sufficiently mastered the material in that chapter. Sentences printed in italics summarise the results of preceding sections. Advanced readers can rapidly scan through the text by watching for these italicised sentences.

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

Goals This book is intended to help its readers to obtain an overview of the rapid developments in quantum information theory, to acquire informed opinions concerning these developments and to understand them with an acceptable investment of time and effort.

Limitations and complementary material The claim to mathematical precision in this book is on the same level as in general textbooks on theoretical physics. Its content is limited to theoretical aspects; a description of the corresponding experiments and technical applications would fill again as many chapters. Each chapter, however, contains a subchapter on complementary topics and more detailed literature; here, references to experiments are included.

These subchapters also contain references to theoretical review articles and books. With their aid, the reader can obtain complementary information about the topics covered in this book and can deepen his or her knowledge. Articles and books with review character were given preference here over the original literature; i.e. the important articles in the sense of the retrospective historical development were not listed consistently, but instead, literature was cited which will be useful to the reader in obtaining a more advanced understanding of the subjects treated.

Content Following the first two chapters, in Chapters 3 and 4, the physics of isolated quantum systems is first developed further. Many examples and applications refer to qubits (2-level-systems). This treatment is then extended through the use of the density operator in Chapter 4. These are the most general states. Chapters 5 and 6 introduce the classical and the quantum-mechanical concepts of entropy and information.

The fundamentals of the physics of composite quantum systems are described in Chapter 7. The fact that an entangled state can contain several subsystems has a number of surprising consequences. An introduction to this topic is given in Chapter 8. Entanglement leads to correlation of the subsystems. The non-locality of the states is accompanied by the possibility of non-local measurements (Chap. 9).

The experimental observation of specific quantum correlations (EPR correlations) confirms the fundamental statement that there is no classical alternative to quantum mechanics (Chap. 10). These EPR correlations can be used as the basis of a quantum cryptography which is in principle completely secure against spying. Quantum teleportation is also based on them (Chap. 12). For the quantum computer, entanglement is an essential tool. Exploitation of quantum parallelism allows a large number of functional values to be computed in only a few operations. The problem is then reading out the results (Chap. 12).

In Chapter 13, we turn to the generalised dynamics of open quantum systems and first discuss generalised measurements, which contain projective measurements as a special case. They play an increasingly important role –together with positive operator-value measures (POVM)– in current publications. The general evolution of open quantum systems between their preparation and their measurement is described with the use of quantum operators. Various quantum channels are discussed in Chap. 14. The generalisations of projective measurements and unitary transformations lead to a new scenario of quantum physics.

Decoherence is the loss of the ability to exhibit interference, and is thus a problem for quantum computers. Conversely, environmentally-induced decoherence plays an important

role in answering the question of why classical objects exist (Chap. 15). It is tempting to apply this approach to obtaining an understanding of quantum measurement processes as well. The book closes with the supplemental proof of several theorems in Chap. 16.

Acknowledgments First of all, I wish to thank my wife for her patience. My cooperation with Thomas Konrad, extending over many years, has contributed extensively to a deepened understanding of this material. Our "Monday meetings" with Thomas Konrad, Michael Nock and Artur Scherer have likewise provided many helpful suggestions, new ideas, and corrections. In particular, our many mutual discussions have had a major influence in maintaining our enthusiasm for the subject. Joseph Demuth accompanied the writing of the manuscript with many helpful remarks and comments. Jan Nötzold and Marcus Kubitzki helped with the preparation of the manuscript, and without the tireless assistance of Stefan Bretzel and in particular of Michael Nock, it would not have been finished in time to meet the deadline. I owe thanks to all of them. Finally, I also wish to thank the Centre for Applied Photonics (CAP) of the University of Konstanz for its support.

Jürgen Audretsch

Konstanz, January 2005

© 2007 WILEY-VCH Verlag GmbH & Co.

1 The Mathematical Framework

It is the goal of quantum theory – just as of every other physical theory – to predict the results of experiments and to justify these predictions. To this end, it is necessary to describe the state of the physical system at the beginning of an experiment. One must also be able to formulate the evolution of the system under external influences and to predict the effect of its interaction with the measurement apparatus. The mathematical framework which has proven most expedient for the formulation of quantum mechanics is the theory of the Hilbert space and probability theory. The fundamental connection between mathematical quantities and physical reality is established by the following associations:

Quantum system	$\leftrightarrow$	Hilbert space.
Quantum state	$\leftrightarrow$	vector in (or, more generally: density operator on) the Hilbert space.
Evolution of the quantum state	$\leftrightarrow$	linear operators, which act on the vectors, or linear operators, which act on the operator space (Liouville space).
Predictions	$\leftrightarrow$	probabilistic statements.

We will describe this basic scheme of the quantum theory in detail. In this chapter, we first collect the required mathematical definitions and theorems. We shall not prove all of the mathematical theorems; in particular, we assume that the reader has already had some contact with quantum theory, so that our treatment here can be brief.

Since we will be concerned exclusively with d-level quantum systems ($d = 2, 3, \ldots$), we make use of a restriction which will greatly simplify our treatment:

> ***General mathematical assumption*:** *We consider only quantum systems which can be described with the aid of a* finite-dimensional *Hilbert space $\mathcal{H}_d$ of dimension $d = 2, 3, \ldots$.*

This restriction is justified, since the essential conceptual problems and the new ideas and central methods can all be introduced by referring to a finite-dimensional Hilbert space. We wish to avoid the addition of mathematical subtleties to the conceptual physical problems. For the majority of physically-relevant cases which require a description in infinite-dimensional Hilbert space, the results for finite-dimensional spaces can be directly applied.

As is usual in theoretical physics, we will make use of the *Dirac notation*. In this framework, it is expedient to place the dyadic decomposition of operators at the centre of our treatment. This is important for practical applications, since it permits a simple, direct reading-off of the properties and effects of the operators.

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

1.1 Hilbert Vector Space

1.1.1 The Scalar Product and the Dirac Notation

A d-dimensional Hilbert space $\mathcal{H}_d$ is a linear complex vector space in which a scalar product is defined. The vectors are denoted by $|\varphi\rangle$, $|\psi\rangle$, $|u\rangle$, $|\Phi\rangle$, etc.; $|\text{null}\rangle$ is the null vector.

Addition and multiplication with a complex number, linear independence, the basis and dimensionality of the Hilbert space $\mathcal{H}_d$ are defined in an analogous way to the corresponding concepts in real vector spaces.

A complex number is associated to a pair of vectors $|\varphi\rangle$ and $|\psi\rangle$ as their *scalar product* or *inner product*, which we write in the form $\langle\varphi|\psi\rangle$. As the basis of this *Dirac notation*[1], we have introduced a *ket space* with the *ket vectors* $|\varphi\rangle, |\psi\rangle, \ldots$ and its dual vector space of the *bra vectors* $\langle\chi|, \langle\theta|, \ldots$ (space of linear functionals). There is a declared correspondence between the vectors of the ket space and of the bra space,

$$|\varphi\rangle \overset{d.c.}{\leftrightarrow} \langle\varphi| \,, \tag{1.1}$$

which is called the *dual correspondence* for vectors. We use the same central symbol as an expression of the dual correspondence. Here, a ket vector $|\varphi\rangle = c_1|\varphi_1\rangle + c_2|\varphi_2\rangle$ is associated via a one-to-one correspondence with a bra vector $\langle\varphi| = c_1^*\langle\varphi_1| + c_2^*\langle\varphi_2|$ ($*$ signifies the complex conjugate). The ordering within the product $\langle\varphi|\psi\rangle$ is thus important. We have:

$$\begin{aligned}
&\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle^* \\
&\langle\varphi|c_1\psi_1 + c_2\psi_2\rangle = c_1\langle\varphi|\psi_1\rangle + c_2\langle\varphi|\psi_2\rangle \,, \quad c_1, c_2 \in \mathbb{C} \\
&\langle\varphi|\varphi\rangle \geq 0 \,\forall\, |\varphi\rangle \in \mathcal{H}_n \,, (\langle\varphi|\varphi\rangle = 0 \Leftrightarrow |\varphi\rangle = |\text{null}\rangle) \,.
\end{aligned} \tag{1.2}$$

From this, it follows that

$$\langle c_1\varphi_1 + c_2\varphi_2|\psi\rangle = c_1^*\langle\varphi_1|\psi\rangle + c_2^*\langle\varphi_2|\psi\rangle \,. \tag{1.3}$$

The scalar product is linear in its second argument and *antilinear* in its first argument. When $\langle\varphi|\psi\rangle = 0$ holds, the vectors are termed *orthogonal* to each other.

The product induces a *norm* on the Hilbert space according to

$$\|\varphi\| =: \||\varphi\rangle\| := \sqrt{\langle\varphi|\varphi\rangle} \,. \tag{1.4}$$

It vanishes if and only if $|\varphi\rangle$ is the zero vector. We mention without proof *Schwarz's inequality*

$$|\langle\varphi|\psi\rangle| \;\leq\; \|\varphi\|\,\|\psi\| \tag{1.5}$$

and the *triangle relations*

$$\|\varphi\| - \|\psi\| \;\leq\; \|\psi - \varphi\| \;, \qquad \|\varphi + \psi\| \;\leq\; \|\varphi\| + \|\psi\| \;. \tag{1.6}$$

[1] Following Dirac, the scalar product is written as $\langle\varphi|\psi\rangle$ and called a "bracket". Its components "bra" $\langle\varphi|$ and "ket" $|\psi\rangle$ denote independent vectors

One can show by substitution that

$$\langle\varphi|\psi\rangle = \frac{1}{4}\left(\|\varphi+\psi\|^2 - \|\varphi-\psi\|^2 + i\,\|\varphi-i\psi\|^2 - i\,\|\varphi+i\psi\|^2\right) \tag{1.7}$$

holds, as well as the *parallelogram equation*

$$\|\varphi+\psi\|^2 + \|\varphi-\psi\|^2 = 2\,\|\varphi\|^2 + 2\,\|\psi\|^2 \ . \tag{1.8}$$

For a set of vectors $\{|\varphi_1\rangle, |\varphi_2\rangle, \ldots, |\varphi_l\rangle\}$ in $\mathcal{H}_d$, span$(|\varphi_1\rangle, \ldots, |\varphi_l\rangle)$ denotes the set of all possible linear combinations of these vectors. This set forms a subspace of $\mathcal{H}_d$ which is itself a Hilbert space. We denote an *orthonormal basis* by *ONB*. For an ONB $\{|i\rangle, i = 1, \ldots, d\}$, the decomposition

$$|\varphi\rangle = \sum_{i=1}^{d} |i\rangle\langle i|\varphi\rangle \tag{1.9}$$

applies, defining the *components* $\langle i|\varphi\rangle$ of the vector $|\varphi\rangle$ with respect to the ONB. The set of all vectors $|\psi\rangle$ which are orthogonal to all the vectors in a subspace $\hat{\mathcal{H}}$ of $\mathcal{H}$ forms an additional subspace of $\mathcal{H}$, which is called the *orthogonal complement* $\hat{\mathcal{H}}^\perp$. The direct sum of the two subspaces is again the Hilbert space, $\mathcal{H} = \hat{\mathcal{H}} \oplus \hat{\mathcal{H}}^\perp := \{\alpha|\chi\rangle + \beta|\psi\rangle$ with $|\chi\rangle \in \hat{\mathcal{H}},\ |\psi\rangle \in \hat{\mathcal{H}}^\perp$ and $\alpha, \beta \in \mathbb{C}\}$.

1.1.2 Linear Operators on the Hilbert Space

Linear operators $A, B, \ldots$ map ket vectors in a linear way onto one another

$$\begin{aligned}
A(\alpha|\psi\rangle + \beta|\phi\rangle) &= \alpha A|\psi\rangle + \beta A|\phi\rangle && \text{linearity}(\alpha, \beta \in \mathbb{C}) \\
(A+B)|\psi\rangle &= A|\psi\rangle + B|\psi\rangle && \text{sum} \\
(AB)|\psi\rangle &= A(B|\psi\rangle) && \text{product} \\
A|\psi_a\rangle &= a|\psi_a\rangle && \textit{eigenvector } |\psi_a\rangle \text{ of } A \\
& && \textit{eigenvalue } a \text{ of } A \\
\mathbb{1}|\psi\rangle &= |\psi\rangle && \textit{identity operator, unit operator}.
\end{aligned} \tag{1.10}$$

The domain of definition of A need not be the entire Hilbert space, and its co-domain need not be identical with its definition range. When necessary, we will make a remark on this point. For the *inverse operator* A^{-1}, we have $AA^{-1} = A^{-1}A = \mathbb{1}$. We wish to extend the Dirac notation further, and therefore adopt the convention that operators on the bra space (arrow to the left) act from the right on bra vectors:

$$\langle\varphi'| = \langle\varphi|\overleftarrow{B}\ . \tag{1.11}$$

The operators on the ket space (arrow to the right) act correspondingly from the left. For the resulting vector, we write

$$|\psi'\rangle = \overrightarrow{A}|\psi\rangle =: |A\psi\rangle\,. \tag{1.12}$$

Via the dual correspondence (1.1), a ket vector $|A\psi\rangle$ corresponds to a bra vector $\langle A\psi|$:

$$|A\psi\rangle \overset{d.c.}{\leftrightarrow} \langle A\psi| \; . \tag{1.13}$$

We in addition introduce a dual correspondence for the operators. Referring to the Dirac notation, the bra operator which corresponds to a ket operator $\overrightarrow{A}$ is likewise denoted by the same central symbol A:

$$\overrightarrow{A} \overset{d.c.}{\leftrightarrow} \overleftarrow{A} \tag{1.14}$$

The correspondence is determined by the following condition on the scalar products (first equation):

$$(\langle\varphi|\overleftarrow{A})|\psi\rangle = \langle\varphi|(\overrightarrow{A}|\psi\rangle) =: \langle\varphi|A|\psi\rangle \; . \tag{1.15}$$

The second equation is an abbreviation, with the compactness which is characteristic of the Dirac notation. Furthermore, we write $A|\psi\rangle$ for $\overrightarrow{A}|\psi\rangle$ and $\langle\psi|A$ for $\langle\psi|\overleftarrow{A}$.

Adjoint operators The dual correspondence leads from $|\psi\rangle$ to $\langle\psi|$. By application of $\overrightarrow{A}$ to $|\psi\rangle$, we obtain $|A\psi\rangle$, and the dual correspondence (1.13) defines $\langle A\psi|$. One can however also obtain $\langle A\psi|$ directly in bra space by applying an operator $\overleftarrow{A}^\dagger$ to $\langle\psi|$:

$$\langle A\psi| =: \langle\psi|\overleftarrow{A}^\dagger \; . \tag{1.16}$$

The operator $\overleftarrow{A}^\dagger$ thus defined is called the *adjoint operator* to $\overleftarrow{A}$. By $\overrightarrow{A}$, both $\overleftarrow{A}$ as well as $\overleftarrow{A}^\dagger$ are defined in the bra space. Finally, via the dual correspondence, the adjoint operator $\overrightarrow{A}^\dagger$ in the ket space is:

$$\overrightarrow{A}^\dagger \overset{d.c.}{\leftrightarrow} \overleftarrow{A}^\dagger \; . \tag{1.17}$$

In the Dirac notation, we can leave off the arrows as in Eq. (1.15) and thereby omit the explicit reference to the two spaces. We evaluate Eq. (1.16) using Eq. (1.15):

$$\langle A\psi|\varphi\rangle = (\langle\psi|\overleftarrow{A}^\dagger)|\varphi\rangle = \langle\psi|(\overrightarrow{A}^\dagger|\varphi\rangle) = \langle\psi|A^\dagger\varphi\rangle = \langle\psi|A^\dagger|\varphi\rangle \tag{1.18}$$

and summarise the result:

$$\langle\overleftarrow{A}\psi|\varphi\rangle = \langle\psi|\overrightarrow{A}^\dagger\varphi\rangle = \langle\psi|A^\dagger|\varphi\rangle \; . \tag{1.19}$$

With $\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle^*$, it follows from Eq. (1.19) that

$$\langle\psi|A^\dagger|\varphi\rangle = \langle\varphi|A\psi\rangle^* = \langle\varphi|A|\psi\rangle^* \; . \tag{1.20}$$

Repeated application of Eq. (1.20) yields

$$\langle\varphi|A|\psi\rangle = (\langle\varphi|A|\psi\rangle^*)^* = \langle\psi|A^\dagger|\varphi\rangle^* = \langle\varphi|(A^\dagger)^\dagger|\psi\rangle \tag{1.21}$$

for arbitrary vectors $\langle\varphi|$ and $\langle\psi|$. Thus, we find

$$(A^\dagger)^\dagger = A \tag{1.22}$$

and we obtain the corresponding relation to Eq. (1.19)

$$\langle A^\dagger\psi|\varphi\rangle = \langle\psi|A\varphi\rangle = \langle\psi|A|\varphi\rangle \; . \tag{1.23}$$

In a similar manner, one can readily convince oneself of the validity of the following operator relations:

$$\left(A^{-1}\right)^\dagger = \left(A^\dagger\right)^{-1} \; , \quad (cA)^\dagger = c^* A^\dagger \tag{1.24}$$

$$(A+B)^\dagger = A^\dagger + B^\dagger \, , \quad (AB)^\dagger = B^\dagger A^\dagger \, . \tag{1.25}$$

In addition to the definition (1.16), equations (1.22) and (1.23) are frequently used.

Dyadic decomposition We introduce the *dyadic product* (outer product) or the *dyad* $|u\rangle\langle v|$ of two vectors $|u\rangle$ and $|v\rangle$. It is a linear operator

$$|\varphi\rangle \rightarrow |\psi\rangle = (|u\rangle\langle v|)|\varphi\rangle = |u\rangle\langle v|\varphi\rangle \, , \tag{1.26}$$

which produces a vector parallel to $|u\rangle$. The scalar product $\langle v|\varphi\rangle$ appears as a complex factor. $|u\rangle\langle v|$ is in the first instance to be considered as an overall symbol which cannot be decomposed, denoting a linear operator with certain properties. These include:

$$(\alpha|u\rangle\langle v|)^\dagger = \alpha^*|v\rangle\langle u| \, . \tag{1.27}$$

For operator products, we find

$$A|u\rangle\langle v| = |Au\rangle\langle v| \, , \qquad\qquad |u\rangle\langle v|A = |u\rangle\langle A^\dagger v| \; . \tag{1.28}$$

As we have seen in Eq. (1.9), the identity operator can be represented in terms of a dyad with the aid of an ONB $\{|i\rangle, i = 1, \ldots, d\}$ of the Hilbert space:

$$\mathbb{1} = \sum_i |i\rangle\langle i| \; . \tag{1.29}$$

This is also referred to as the *completeness relation* or the *dyadic decomposition of the identity operator*. *From Eq. (1.29), it follows using Eq. (1.26) that every linear operator has a* dyadic decomposition *(outer product representation)*

$$A = \sum_{i,j} |i\rangle\langle i|A|j\rangle\langle j| = \sum_{i,j} \langle i|A|j\rangle|i\rangle\langle j| = \sum_{i,j} A_{ij}|i\rangle\langle j| \tag{1.30}$$

with the matrix elements $A_{ij} := \langle i|A|j\rangle$. We can read off the equations (1.26) through (1.30) as a suggestive mnemonic rule, *that* $|u\rangle\langle v|$ *and the dyadic decomposition of* A *can be taken formally to act in such a way, as if* $|u\rangle$ *and* $\langle v|$ *were independent vectors and not parts of*

an overall symbol $|u\rangle\langle v|$. This is one of the great advantages of the Dirac notation. For the adjoint operator, we obtain

$$A^\dagger = \sum_{i,j} A^*_{ij} |j\rangle\langle i| \; . \tag{1.31}$$

Via the *supremum norm* $\|A\|$, one can associate to a linear operator A a positive number:

$$\|A\| := \sup_{\langle\varphi|\varphi\rangle=1} |\langle\varphi|A|\varphi\rangle| \, . \tag{1.32}$$

Trace The *trace* is a frequently-used complex-valued function of a linear operator:

$$\mathrm{tr}[A] := \sum_i \langle i|A|i\rangle = \sum_i A_{ii} \, , \quad \{|i\rangle\} \text{ ONB} \; . \tag{1.33}$$

The trace of an operator is independent of the choice of the basis. The proof of this statement demonstrates the usefulness of the dyadic decomposition (1.29) of the identity operator. Let $\{|l_i\rangle\}$ and $\{|m_j\rangle\}$ be an arbitrary ONB; then using the mnemonic rule above, we find:

$$\begin{aligned} \mathrm{tr}[A] &= \sum_i \langle l_i|A|l_i\rangle = \sum_{i,j,k} \langle l_i|m_j\rangle\langle m_j|A|m_k\rangle\langle m_k|l_i\rangle \\ &= \sum_{i,j,k} \langle m_k|l_i\rangle\langle l_i|m_j\rangle\langle m_j|A|m_k\rangle = \sum_{j,k} \langle m_k|m_j\rangle\langle m_j|A|m_k\rangle \\ &= \sum_j \langle m_j|A|m_j\rangle \; . \end{aligned} \tag{1.34}$$

In a similar manner, using Eq. (1.29) one can prove the following properties of the trace:

$$\begin{aligned} \mathrm{tr}[AB] &= \mathrm{tr}[BA] && \text{cyclic permutations} \\ \mathrm{tr}[A+B] &= \mathrm{tr}[A] + \mathrm{tr}[B] && \text{linearity} \\ \mathrm{tr}[\alpha A] &= \alpha \, \mathrm{tr}[A] && \text{linearity} \\ \mathrm{tr}[A|\psi\rangle\langle\psi|] &= \langle\psi|A|\psi\rangle && \text{expectation value of } A \\ \mathrm{tr}[|\varphi\rangle\langle\psi|] &= \langle\psi|\varphi\rangle && \text{trace of a dyad} \\ \mathrm{tr}[A^\dagger] &= (\mathrm{tr}[A])^* && \text{adjoint operator} \end{aligned} \tag{1.35}$$

The terms *expectation value* or *mean value* of A used in quantum physics will later be justified on the basis of physical arguments.

1.1.3 Normal Operators and Spectral Decompositions

Among the linear operators on $\mathcal{H}_d$, those which are diagonalisable, also called the *normal operators*, play a particularly important role in mathematics and physics. An operator N is termed *diagonalisable* if there exists an ONB $\{|i\rangle\}$ of $\mathcal{H}_d$ and a set of complex numbers $\lambda_i \in \mathbb{C}$ such that

$$N|i\rangle = \lambda_i |i\rangle \tag{1.36}$$

holds. Here, $\lambda_i = 0$ is not excluded. An immediate result is that the matrix of N in the ONB of the eigenvectors is diagonal

$$N_{ij} = \langle i|N|j\rangle = \lambda_i \delta_{ij} \tag{1.37}$$

and therefore the operator N can be written in the form of a *spectral decomposition*

$$N = \sum_i \lambda_i |i\rangle\langle i| \,, \quad \lambda_i \in \mathbb{C} \,. \tag{1.38}$$

This is also called the *orthogonal decomposition*. The ONB $\{|i\rangle\}$ of Eq. (1.36) is also referred to as the *eigenbasis* of N. Conversely, the diagonalisability condition (1.36) follows directly from each of these relations. If there are $g \geq 2$ linearly-independent eigenvectors $|j_l\rangle$ belonging to an eigenvalue λ_j of the eigenvalue problem (1.36), where l = 1... g, then λ_j is said to be *g-fold degenerate*. Every linear combination of these eigenvectors

$$|\psi\rangle = \sum_{l=1}^{g} c_l |j_l\rangle \tag{1.39}$$

is then likewise an eigenvector belonging to the eigenvalue λ_j. The eigenvectors span a g-dimensional subspace $\mathcal{H}_{(j)}$ of $\mathcal{H}$. The *projector*

$$P = \sum_{l=1}^{g} |j_l\rangle\langle j_l| \,, \quad P^\dagger = P \,; \quad P^2 = P \,, \tag{1.40}$$

projects into the subspace $\mathcal{H}_{(j)}$. The projector $Q = 1 - P$ projects into the orthogonal complement of $\mathcal{H}_{(j)}$, i.e. $\mathcal{H}_{(j)}^{\perp}$. The subspaces belonging to different eigenvalues are orthogonal to one another.

Diagonalisability is by no means a trivially-occurring property. Even in the two-dimensional Hilbert space $\mathcal{H}_2$, there are frequently-used operators which are not diagonalisable. An example is

$$A = |0\rangle\langle 1| \quad \text{with} \quad \langle 0|1\rangle = 0 \quad \text{and} \quad \langle 0|0\rangle = \langle 1|1\rangle = 1 \tag{1.41}$$

as can be shown with the help of the following theorem.

In order to recognise whether a given operator is a normal operator, the following central theorem is very useful: *A necessary and sufficient condition that an operator N can be spectrally decomposed – that is, it is diagonalisable – is the vanishing of the commutator $([A, B]_- := AB - BA)$ of N and $N^\dagger$:*

$$[N, N^\dagger]_- = 0 \,. \tag{1.42}$$

The proof of this theorem can serve as an example of the application of the formalism which we have thus far constructed. The fact that diagonalisability follows from Eq. (1.42) is clear. The converse direction of the proof can be divided into two steps:

1. Step: Each operator in $\mathcal{H}_d$ has at least one eigenvalue λ and one eigenvector $|1\rangle$, which can be found with the aid of the secular equation:

$$N|1\rangle = \lambda|1\rangle\,, \quad \langle 1|N^\dagger = \lambda^*\langle 1|\,. \tag{1.43}$$

It then follows that

$$\langle 1|N|1\rangle = \lambda\,, \quad \langle 1|N^\dagger|1\rangle = \lambda^* \tag{1.44}$$

and thus

$$N^\dagger|1\rangle = \lambda^*|1\rangle + |a\rangle\,, \quad \langle 1|N = \lambda\langle 1| + \langle a| \tag{1.45}$$

with $\langle 1|a\rangle = 0$. Using the normality condition $[N, N^\dagger]_- = 0$, we find after evaluation with Eqs. (1.43) and (1.45)

$$0 = \langle 1|[N, N^\dagger]_-|1\rangle = \langle a|a\rangle\,. \tag{1.46}$$

$|a\rangle$ is thus the null vector $|\text{null}\rangle$ and (1.45) can be written as follows:

$$N^\dagger|1\rangle = \lambda^*|1\rangle\,, \quad \langle 1|N = \lambda\langle 1|\,. \tag{1.47}$$

We have thus determined the action of N and $N^\dagger$ on $|1\rangle$.

2. Step: We complete $|1\rangle$ to obtain an ONB $\{|i\rangle\}$ and introduce with the aid of the dual notation for N:

$$N = \sum_{ij} n_{ij}|i\rangle\langle j|, \quad n_{ij} := \langle i|N|j\rangle, \quad n_{1i} = n_{i1} = \lambda\delta_{i1} \tag{1.48}$$

the operator M

$$M := N - \lambda|1\rangle\langle 1|\,, \quad M = \sum_{i,j\neq 1} n_{ij}|i\rangle\langle j|\,. \tag{1.49}$$

M is the restriction of N to the orthogonal complement of $|1\rangle$.

Making use of Eqs. (1.43) and (1.47), we can show that M is also a normal operator, $([M, M^\dagger]_- = 0)$. The same procedure can be applied to it in the subspace which is orthogonal to $|1\rangle$. M also has an eigenvector, which we denote as $|2\rangle$. We now complete $|1\rangle$ and $|2\rangle$ to an ONB and repeat the procedure. We continue in the same manner until the entire Hilbert space is used up and $|1\rangle$ has been completed to a well-defined ONB. At the same time, N has been spectrally decomposed with respect to this basis. This concludes the proof. The fact that the operator A of Eq. (1.41) does *not* fulfill the condition (1.42) can be readily verified.

The diagram in Fig. 1.1 demonstrates how the various properties of the operators in Hilbert space correspond in an intuitively clear way to an increasing specialisation in the dyadic decomposition. In the following section, we will go through this diagram step by step from above to below.

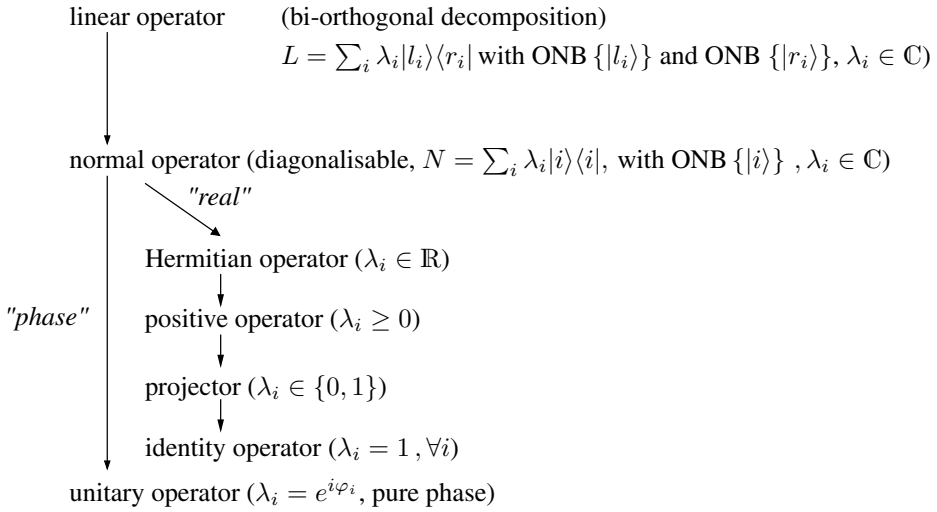

linear operator
(bi-orthogonal decomposition)
$L = \sum_i \lambda_i |l_i\rangle\langle r_i|$ with ONB $\{|l_i\rangle\}$ and ONB $\{|r_i\rangle\}$, $\lambda_i \in \mathbb{C}$)
normal operator (diagonalisable, $N = \sum_i \lambda_i |i\rangle\langle i|$, with ONB $\{|i\rangle\}$, $\lambda_i \in \mathbb{C}$)
"real"
Hermitian operator ($\lambda_i \in \mathbb{R}$)
"phase"
positive operator ($\lambda_i \geq 0$)
projector ($\lambda_i \in \{0, 1\}$)
identity operator ($\lambda_i = 1\,, \forall i$)
unitary operator ($\lambda_i = e^{i\varphi_i}$, pure phase)

Figure 1.1: Hierarchy of operators. Characterisation of operators through their dyadic decomposition. $\rightarrow$ is in each case the direction of increasing specialisation. The eigenvalues are characterised in brackets ( ). The bi-orthogonal decomposition of a linear operator is derived in Sect. 13.3.3.

Functions of operators An *operator function* $f(N)$ is defined in terms of its expansion in a power series. For a normal operator N in the dyadic decomposition, it can be expressed in a simple way in terms of functions of the eigenvalues:

$$f(N) := \sum_i f(\lambda_i)|i\rangle\langle i| \quad \Rightarrow \quad f(N)|i\rangle = f(\lambda_i)|i\rangle \; . \tag{1.50}$$

$f(N)$ has the same eigenvectors $|i\rangle$ as N. We give an example which is formulated as a matrix representation with respect to the basis of the eigenvectors:

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = |0\rangle\langle 0| - |1\rangle\langle 1| \tag{1.51}$$

$$e^{\varphi\sigma_z} = e^{\varphi}|0\rangle\langle 0| + e^{-\varphi}|1\rangle\langle 1| = \begin{pmatrix} e^{\varphi} & 0 \\ 0 & e^{-\varphi} \end{pmatrix} \; . \tag{1.52}$$

1.1.4 Hermitian Operators

We follow the right-hand branch of the tree diagram in Fig. 1.1. A linear operator H on $\mathcal{H}_d$ is termed *Hermitian* or *self-adjoint* when it has the property $H^\dagger = H$. *Hermitian operators are special normal operators*. They play an important role in quantum mechanics, owing to their special properties: *Hermitian operators have a spectral decomposition with an ONB* $\{|i\rangle\}$

$$H = \sum_i r_i |i\rangle\langle i| \, , \quad r_i \in \mathbb{R} \tag{1.53}$$

and real eigenvalues r_i. In case of degeneracy, the eigenvectors can be chosen to be orthonormal, so that $\{|i\rangle\}$ forms an ONB. Eigenvectors belonging to different eigenvalues are orthogonal. This is often called the *spectral theorem*. Hermitian operators are also referred to as *observables*. The reason for this physical terminology will become clear later.

It follows immediately from Eq. (1.53) together with Eq. (1.35) that for an arbitrary vector $|\varphi\rangle$, the *expectation value* $\langle\varphi|H|\varphi\rangle$ is real. It is an important characteristic of Hermitian operators that the converse is also true: *the expectation value* $\langle\varphi|A|\varphi\rangle$ *is real for all vectors if and only if A is Hermitian.*

For a proof of the converse, we assume that for an operator A the mean value $\langle\chi|A|\chi\rangle$ is real for all vectors $|\chi\rangle$. For two arbitrary vectors $|\varphi\rangle$ and $|\psi\rangle$ from $\mathcal{H}$, the identity

$$\begin{aligned} 4\langle\varphi|A|\psi\rangle = \quad & \{(\langle\varphi| + \langle\psi|)A(|\varphi\rangle + |\psi\rangle) - (\langle\varphi| - \langle\psi|)A(|\varphi\rangle - |\psi\rangle)\} \\ + \quad & i[(\langle\varphi| + i\langle\psi|)A(|\varphi\rangle - i|\psi\rangle) - (\langle\varphi| - i\langle\psi|)A(|\varphi\rangle + i|\psi\rangle)] \end{aligned} \tag{1.54}$$

holds. If we exchange $|\varphi\rangle$ and $|\psi\rangle$ in this expression, then the part denoted by $\{\dots\}$ remains the same and the part $[\dots]$ changes its sign. Taking into account that all expectation values[2] are real, it then follows that $\langle\psi|A\varphi\rangle = \langle\varphi|A\psi\rangle^* = \langle A\psi|\varphi\rangle$. The operator A is thus Hermitian. It is notable that Eq. (1.54) contains on the right only expectation values, and on the left only a transition matrix element. *When all the expectation values of an Hermitian operator are known, then all the transition matrix elements are also known.*

The expectation value $\langle\varphi|A|\varphi\rangle$ is also called the *mean value*. Since their eigenvalues and mean values are real, Hermitian operators will play a special role in the theory of measurements (cf. Chap. 2).

Commuting Hermitian operators For these, the theorem on simultaneous diagonalisability holds (w/o.P.)[3]: *Two Hermitian operators (observables)* A *and* B *commute (*$[A, B]_- = 0$*) if and only if they have a common ONB* $\{|i\rangle\}$ *of eigenvectors.*

If the eigenvalue a of an observable A is degenerate, then the eigenvectors form a subspace which is at least two-dimensional. No associated eigenvector is therefore uniquely characterised by specifying a. If we consider only those eigenvectors of A within the subspace which are at the same time eigenvectors of an observable B which commutes with A and has eigenvalues b (intersecting sets), then a common eigenvector could be uniquely specified through this additional condition. We denote it by $|a, b\rangle$:

$$A|a, b\rangle = a|a, b\rangle, \quad B|a, b\rangle = b|a, b\rangle \; . \tag{1.55}$$

If only a subspace is determined in this way, then we continue and require that an eigenvector of A and B at the same time be an eigenvector of an observable C which commutes with A and B: $|a, b, c\rangle$. This procedure must be repeated until all degeneracies have been lifted. A set of observables which possesses exactly one common system of eigenvectors is called a *complete system of commuting observables*. Specifying the eigenvalues of all the operators determines a vector precisely. It is important that the procedure described in fact terminates. This is guaranteed by the following result (w/o.P.): *In every Hilbert space* $\mathcal{H}_\mathrm{d}$*, there exists a finite(!) complete set of operators which commute pairwise (functions of operators are not taken into consideration).* For the proof, we refer to the literature (cf. Sect. 1.4).

[2] According to Section 1.1.1, $\langle\varphi| + i\langle\psi|$ is the dual bra vector of $|\varphi\rangle - i|\psi\rangle$.

[3] w/o.P. means *without proof*

1.1.5 Unitary Operators

We first follow the left-hand branch of the operator-hierarchy tree in Fig. 1.1 and thereafter return to the right-hand branch. A linear operator U is called *unitary* when it has the property $U^\dagger = U^{-1}$. *Unitary operators are special normal operators. They have a spectral decomposition*

$$U = \sum_i e^{i\varphi_i} |i\rangle\langle i| \,, \quad \varphi_i \in \mathbb{R}, \tag{1.56}$$

with an ONB $\{|i\rangle\}$*, whereby, due to the defining equation, the eigenvalues are pure "phase factors". As with Hermitian operators, the eigenvectors span the entire space.* Eigenvectors with different eigenvalues are orthogonal. Eigenvectors with degenerate eigenvalues can be chosen to be orthogonal. As one can readily show, a linear operator is unitary precisely when each of its matrix representations is unitary. It follows immediately from the spectral decomposition that the operator function $U(t) = e^{iHt}$, $t \in \mathbb{R}$, is unitary if H is Hermitian. Furthermore, in this case:

$$U(t=0) = \mathbb{1} \tag{1.57}$$

$$U(t_2)U(t_1) = U(t_2 + t_1) \,. \tag{1.58}$$

Unitary equivalence and conservation of the norm Under combined unitary transformations of vectors and operators according to

$$|\varphi'\rangle = U|\varphi\rangle \quad A' = UAU^{-1} \,, \tag{1.59}$$

scalar products (in particular, the norm of a vector), eigenvalues and expectation values remain unchanged. *Conversely, a linear operator T, which conserves the norm on application to an arbitrary vector in $\mathcal{H}_d$,*

$$\|T\varphi\| = \|\varphi\| \,, \tag{1.60}$$

is a unitary operator: $T^\dagger = T^{-1}$. For the proof, we apply Eq. (1.7) and rewrite it using Eq. (1.60). For T, a unitarity relation holds:

$$\langle T\varphi|T\psi\rangle = \langle\varphi|\psi\rangle \,. \tag{1.61}$$

1.1.6 Positive Operators and Projection Operators

We wish to discuss some special cases of Hermitian operators (compare Fig. 1.1). A *positive operator* is defined by the fact that for an arbitrary vector $|\varphi\rangle$, the following inequality:

$$\langle\varphi|A|\varphi\rangle \geq 0 \quad \forall\, |\varphi\rangle \,, \tag{1.62}$$

holds, i.e. its expectation value is always real and non-negative. We can then write

$$A \geq 0 \,. \tag{1.63}$$

Furthermore, we define an inequality for operators:

$$A \geq B \Leftrightarrow (A - B) \geq 0 \,. \tag{1.64}$$

From the condition of positivity, it follows for the spectral decomposition that *every positive operator A is Hermitian, $A^\dagger = A$. It has the spectral decomposition*

$$A = \sum_i a_i |i\rangle\langle i|, \quad a_i \geq 0 \,. \tag{1.65}$$

with non-negative eigenvalues.

For an arbitrary linear operator A, $A^\dagger A$ is a positive operator. On the other hand, for each positive operator A there is a linear operator B such that A can be written in the form

$$A = B^\dagger B \,. \tag{1.66}$$

B is determined only up to unitary transformations ($B \to UB$). We can find B explicitly via the spectral decomposition of A (1.65) and an ONB $\{|\varphi_i\rangle\}$

$$B = \sum_i \sqrt{a_i}\, |\varphi_i\rangle\langle i| \,. \tag{1.67}$$

Substitution verifies (1.66).

A linear operator P is a *projection operator* (more precisely: an orthogonal projection operator) when it meets the following conditions:

$$P^2 = P \quad \textit{idempotent.} \tag{1.68}$$

$$P^\dagger = P \quad \text{Hermitian.} \tag{1.69}$$

It follows from these properties that

$$\langle v|P|v\rangle = \langle v|PP|v\rangle = \langle v|P^\dagger P|v\rangle = \|P|v\rangle\|^2 \geq 0 \,. \tag{1.70}$$

P is therefore a positive operator and fulfills

$$P = \sum_i p_i |i\rangle\langle i| \,; \quad p_i \geq 0 \tag{1.71}$$

with the ONB $\{|i\rangle\}$. Because it is idempotent, we furthermore have

$$P^2 = \sum_i p_i^2 |i\rangle\langle i| \,, \quad P = \sum_i p_i |i\rangle\langle i| \,, \tag{1.72}$$

and thus $p_i^2 = p_i$ or $p_i \in \{0, 1\}$. The projection operator P therefore assumes the form

$$P = \sum_{j \in I} |j\rangle\langle j|, \qquad I \leftrightarrow \text{subset of the ONB} \,. \tag{1.73}$$

P projects onto the subspace spanned by $\{|j\rangle\}$ with $j \in I$.

As a complement to Fig. 1.1, Fig. 1.2 shows retrospectively the "intersecting sets" of the different types of operators.

1.2 Liouville Operator Space

As we shall see in Chap. 2, in the special case of pure states, quantum-mechanical systems can be described by normalised vectors $|\psi\rangle$ in a Hilbert space $\mathcal{H}_{\rm d}$. In the general case of mixed quantum states, their description is accomplished using density operators (Chap. 4). All possible dynamical changes of state can be described in terms of linear transformations between density operators (Schrödinger representation). We will discuss this quite generally in Chap. 14. In preparation for this discussion, it is expedient to introduce the Liouville space $\mathbb{L}$ here. It is the space of the linear operators which act on the Hilbert space. We can restrict ourselves to a brief presentation, since the procedure is essentially a repetition of Sect. 1.1.

1.2.1 Scalar Product

The *Liouville space* $\mathbb{L}$ is a linear complex vector space whose elements $|A), |B), \ldots$ are the linear operators $A, B, \ldots$ which operate on a Hilbert space. One can readily verify that these linear operators in fact fulfill the axioms of a linear vector space. Later, we will leave off the brackets $|)$ in order to simplify our notation.

In this new notation, the dyadic decomposition (1.30) of an operator A in terms of the basis $\{|i\rangle\}$ on $\mathcal{H}_d$ has the form

$$|A) = \sum_{i,j=1}^{d} A_{ij} ||i\rangle\langle j|) \; . \tag{1.74}$$

The d^2 dyads $|i\rangle\langle j|$ in $\mathcal{H}_d$ make up the d^2 elements $||i\rangle\langle j|)$ of a basis of $\mathbb{L}$. For the dimensions of the spaces, we therefore have

$$\dim \mathbb{L} = (\dim \mathcal{H}_d)^2 \; . \tag{1.75}$$

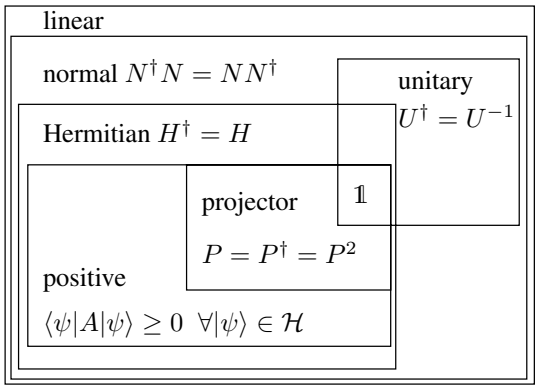


Figure 1.2: "Intersecting sets" of operator types. Note that for $\lambda_i \in \{1, -1\}$, special Hermitian operators can also be unitary and *vice versa*.

Of course, there are other basis sets besides the dyads in $\mathbb{L}$. The Liouville space $\mathbb{L}$ has a *scalar product* $(A|B)$. Formally, it has the same properties as the scalar product in the Hilbert space $\mathcal{H}_d$ (cf. Sect. 1.1.1). $(A|B)$ is a complex number and obeys the relations:

$$(A|B) = (B|A)^* \ , \ (A|c_1B_1 + c_2B_2) = c_1(A|B_1) + c_2(A|B_2) \ , \ (A|A) \geq 0 \ . \quad (1.76)$$

The operator basis Two operators A and B are called *orthogonal* if

$$(A|B) = 0 \quad (1.77)$$

holds, without either of the operators being the null operator. The triangle inequality (1.6) and an equation analogous to the parallelogram equation (1.8) also apply. Every operator $|A)$ can be decomposed in terms of an *orthonormal operator basis* $\{|Q_s), s = 1, \ldots, d^2\}$ of $\mathbb{L}$

$$(Q_s|Q_t) = \delta_{st}, \quad \sum_{s=1}^{d^2} |Q_s)(Q_s| = \mathbb{1} \ : \quad (1.78)$$

$$|A) = \sum_{s=1}^{d^2} |Q_s)(Q_s|A) \ . \quad (1.79)$$

The scalar product as trace Scalar products in $\mathbb{L}$ can be represented in various ways. *We will use the scalar product defined via the trace in $\mathcal{H}_d$*:

$$(A|B) := \mathrm{tr}[A^\dagger B] \ , \quad (1.80)$$

since in this case the Pauli spin operators, which are important for the simplest quantum systems, can be completed to form a basis (compare Sect. 3.1). The decomposition (1.79) –leaving off the vector brackets– takes on the form

$$A = \sum_{s=1}^{d^2} Q_s \, \mathrm{tr}[Q_s^\dagger A] \ . \quad (1.81)$$

The basis of the Liouville space generated from the dyads $|i\rangle\langle j|$ with $i, j = 1, \ldots, d$ is orthonormal with respect to the trace-scalar product (1.80)

$$\Big(|i\rangle\langle j| \Big| |i'\rangle\langle j'|\Big) = \delta_{ii'}\delta_{jj'} \ . \quad (1.82)$$

1.2.2 Superoperators

As might be presumed, we can define *linear* operators in the Liouville space itself, which map the elements of the space onto one another:

$$|A) \to \mathcal{S}|A) =: |\mathcal{S}A) =: \mathcal{S}(A) =: \mathcal{S}A \ . \quad (1.83)$$

These operators, which we write using italic symbols, are called *superoperators*. From the point of view of the Hilbert space $\mathcal{H}_d$, they map linear operators in a linear manner onto one another

$$A \to B = \mathcal{S}(A) \ . \quad (1.84)$$

Examples We give two examples of superoperators: For the superoperator $\mathcal{A}$

$$B \to \mathcal{A}(B) := ABA^{-1} \,, \tag{1.85}$$

linearity follows from the linearity of A. One can readily verify that

$$\mathcal{A}^{-1}(B) = A^{-1}BA \tag{1.86}$$

holds. An important superoperator for the description of the dynamic evolution of mixed states (compare Chap. 4) is the *Liouville operator* or *Liouvillian*, $\mathcal{L}$

$$A \to \mathcal{L}(A) := \frac{1}{\hbar}\,[H, A]_- \;. \tag{1.87}$$

In its application to physical problems, H in this expression is the Hamiltonian. The powers of $\mathcal{L}$ are written

$$\mathcal{L}^2(A) = \frac{1}{\hbar^2}\,\big[H, [H, A]_-\big]_- \;. \tag{1.88}$$

The concepts of adjoint, Hermitian, unitary and positive superoperators can be directly taken over from the corresponding definitions in Hilbert space.

1.3 The Elements of Probability Theory

As we have already emphasized, it is the central goal of quantum theory to make predictions concerning the probability of occurrence of measured values. To this end, it will be assumed that information about the state of the quantum object being measured is available. With this goal in mind, it is expedient to review briefly the basic concepts of probability theory here.

Predictions are conclusions drawn from the past and applied to the future. In classical physics, the reverse direction of conclusions plays a similarly important role. From the results of measurements, conclusions about the state of the object before the measurement are drawn. To what extent is this also possible for quantum systems? In the discussion of this question, Bayes' Theorem plays a very important role. We will sketch its proof after first presenting some preliminary considerations concerning conditional probabilities.

1.3.1 The Probability of Random Events

When a stochastic experiment is repeated, the result cannot be predicted. It is a *random event*. Such events could be for example the occurrence of an even or an odd number of dots when throwing dice, or the occurrence of a number greater than 2. Let $\{A_i;\, i = 1, \ldots, n\}$ be the number of such events. We introduce the following notation, in analogy to set theory:

$A_i \cap A_j \cap A_k$ is the event which consists of the simultaneous occurrence of the events A_i, A_j and A_k. In the case of dice, A_1 could be e.g. the event "even number of dots" and A_2 the event "number of dots > 4"; then $A_1 \cap A_2$ is the event "the six is thrown". $p(A_1 \cap A_2)$ is the probability that *both* A_1 *and also* A_2 occur (*joint probability*). We can also write

$$p(A_1, A_2) := p(A_1 \cap A_2) \,. \tag{1.89}$$

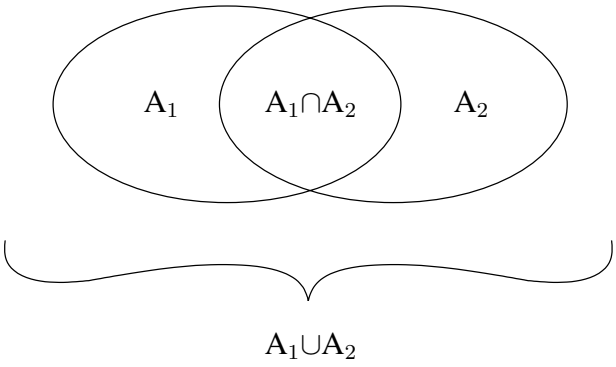


Figure 1.3: Set diagram for probabilities.

$A_i \cup A_j \cup A_k$ is the event consisting of the occurrence of *at least one* of the events A_i, A_j or A_k. For a number Z of dots, let $2 \leq Z \leq 4$ be the event A_1 and $3 \leq Z \leq 5$ be the event A_2. Then $A_1 \cup A_2$ is the event $2 \leq Z \leq 5$.

An impossible event is denoted by $\varnothing$ and a certain event by Ω. Two events A_i and A_j are called *exclusive events* when $A_i \cap A_j = \varnothing$. They cannot occur simultaneously.

Axioms With each random event A we associate a real number $p(A)$ with $0 \leq p(A) \leq 1$, which is called the *probability* of A, and which fulfills a series of axioms that we shall not list here. An example is given by Kolmogorov's axioms. We note only the *additivity axiom*: for pairwise exclusive random events $A_1, A_2, \ldots, A_n$ (i.e. $\mathrm{tr}(A_i \cap A_j) = 0$),

$$p(A_1 \cup A_2 \cup \ldots \cup A_n) = p(A_1) + p(A_2) + \ldots + p(A_n) \tag{1.90}$$

holds. When the events A_1 and A_2 are not exclusive, we find

$$p(A_1 \cup A_2) = p(A_1) + p(A_2) - p(A_1 \cap A_2)\ . \tag{1.91}$$

The set diagram in Fig. 1.3 gives an intuitive picture of this relation. For thrown dice, let $Z \leq 2$ be event A_1 and $Z \geq 4$ be event A_2; then the probability that either A_1 or A_2 occurs is $p(A_1 \cup A_2) = \frac{2}{6} + \frac{3}{6} = \frac{5}{6}$.

Frequency interpretation In order to make the axiom clear, we used the example of throwing dice. In fact, this axiom, like all mathematical axioms, requires no physical interpretation. $p(A)$ is defined by the axioms themselves. When applied to physical events, probability is usually interpreted as the *relative frequency*:

$$p(A) := \lim_{N \to \infty} \frac{N(A)}{N} \tag{1.92}$$

Here, $N(A)$ is the absolute frequency of occurrence of the event A in a total number N of attempts. This physical interpretation is not without problems. For a finite number N, it can be taken as an estimate of $p(A)$.

1.3.2 Conditional Probability and Bayes' Theorem

We extend the concept of probability. The *conditional probability* $p(A|B)$ of an event A is the probability of occurrence of A *under the condition that* another event B, which itself has the probability $p(B)$, has already occurred. We define:

$$p(A|B) := \frac{p(A \cap B)}{p(B)} \ . \tag{1.93}$$

Resolution of this expression leads to the plausible equation for the probability $p(A \cap B)$ for the occurrence of both A and B:

$$p(A \cap B) = p(A|B) \cdot p(B) \ . \tag{1.94}$$

As an example, we consider two urns. The urn U_1 contains 3 white and 3 black balls; urn U_2 contains 2 white and 4 black balls. From each of the urns, balls are picked with the same probability $p(U_1) = p(U_2) = \frac{1}{2}$. The probability of being picked is the same for every ball, i.e. $\frac{1}{12}$. The probability of picking a ball both from U_1 and also that the ball picked be white is given by $p(w \cap U_1) = \frac{3}{12} = \frac{1}{4}$. The conditional probability $p(w|U_1)$ of getting a white ball when picking from urn U_1 is given according to Eq. (1.93) by

$$p(w|U_1) = \frac{p(w \cap U_1)}{p(U_1)} = \frac{2}{4} = \frac{1}{2} \ . \tag{1.95}$$

This follows intuitively directly from the description of the randomness of the situation. Analogously, one finds $p(w|U_2) = \frac{1}{3}$.

Independence Two random events A and B are called *independent* events when the occurrence of the one event has no influence on the probability of occurrence of the other,

$$p(A|B) = p(A) \ . \tag{1.96}$$

In this case, it follows with (1.93) that

$$p(A \cap B) = p(A)p(B) \ . \tag{1.97}$$

From this, it must be distinguished whether the events A and B are exclusive (mutually contradictory), $A \cap B = \varnothing$. In that case, we have $p(A|B) = 0$.

Total probability The certain event Ω can be represented as the sum of n pairwise exclusive random events A_i, $(A_i \cap A_j = \varnothing, \quad \forall i \neq j)$:

$$\Omega = A_1 \cup A_2 \cup \ldots \cup A_n; \ A_i \cap A_j = \varnothing, \quad \forall \, i \neq j \ . \tag{1.98}$$

For an arbitrary random event B, we then find $B = (A_1 \cap B) \cup (A_2 \cap B) \cup \ldots \cup (A_n \cap B)$. From the additivity axiom (1.90) it follows that

$$p(B) = \sum_{i=1}^{n} p(B \cap A_i) \ , \tag{1.99}$$

and with Eq. (1.94) we obtain the *addition theorem for probabilities* or the total probability

$$p(B) = \sum_{i=1}^{n} p(B|A_i)p(A_i) \ . \tag{1.100}$$

We will give an example in the next section.

Bayes' theorem With $p(A \cap B) = p(B \cap A)$, Eq. (1.94) for random events A_i leads to

$$p(A|B)p(B) = p(B|A)p(A) \ . \tag{1.101}$$

Under the assumption of pairwise exclusivity and completeness (1.98), we obtain with Eq. (1.100) the fundamental Bayes' Theorem

$$p(A_i|B) = \frac{p(B|A_i)p(A_i)}{\sum_{j=1}^{n} p(B|A_j)p(A_j)} \ . \tag{1.102}$$

The denominator guarantees the normalisation $\sum_i p(A_i|B) = 1$, which ensures that one of the events A_i must occur.

Bayes' theorem can be interpreted as follows: let the probabilities $p(A_i)$ and the conditional probabilities $p(B|A_i)$ for B given A_i be known for a certain situation. Then equation (1.102) permits the computation of the conditional probability $p(A_i|B)$ for A_i given B. If the event B occurs *after* event A_i, then $p(A_i|B)$ answers the following question: if B occurs, what was the probability that A_i had already occurred? This conclusion for A_i given B applies in the reverse direction to that of $p(B|A_i)$. This demonstrates the significance of the theorem.

We give an example, again based on picking balls from urns. Let us presume the existence of three urns of type I with 2 white and 6 black balls in each one, and of one urn of type II with 1 white and 7 black balls. The probability of choosing any one of the urns from which to pick a ball is the same. Result B means that a white ball is picked. The event A_1 (or A_2) means that a ball is picked from an urn of type I (or of type II). Then we have the following probabilities: $p(A_1) = \frac{3}{4}, p(A_2) = \frac{1}{4}, p(B|A_1) = \frac{1}{4}, p(B|A_2) = \frac{1}{8}$. The probability that a white ball picked comes from an urn of type I is, according to Bayes' theorem, given by $p(A_1|B) = \frac{6}{7} = 0.86$ and is thus greater than $p(A_1)$. A white ball comes from the urn of type II with a probability $p(A_2|B) = \frac{1}{7} = 0.14$, which is less than $p(A_2)$. The choice of the type of urn is made with the *a priori* probabilities $p(A_i)$. If a white ball was picked, one can make an inference about which urn it originated from. For this *inference* there is in general only a probability statement which is given by $p(A_i|B)$. If the urn of type II contained no white balls, the inference could be made with certainty $(p(A_1|B) = 1)$ that a white ball was picked from an urn of type I.

The following explanation of Bayes' theorem can also be helpful: we consider the special case that all the $p(A_i)$ are equal. The event B is to be predicted. The event A_k for which the subsequent occurrence of B is most probable (i.e. $p(B|A_k) = \max$) has also previously occurred with maximum probability, $p(A_k|B) = \max$.

Bayes' assumption This is not to be confused with Bayes' theorem. If there is no reason to assume that an event A_i is particularly preferred by the situation, it can be reasonable to make *Bayes' assumption* that all the *a priori* probabilities are equal,

$$p(A_1) = p(A_2) = \ldots = p(A_n) \; . \tag{1.103}$$

This assumption is not the same as Bayes' theorem. After the occurrence of B, this assumption is replaced by the probabilities $p(A_i|B)$ from Eq. (1.102). The probabilities can be estimated in this way.

1.3.3 Random Quantities

A *random variable* X is given by association of numbers x to the corresponding random events. Throws of dice can be used as an illustrative example. A discrete random quantity X is determined by the values $x_1, x_2, \ldots, x_n$ and the probabilities $p(x_1), p(x_2), \ldots, p(x_n)$ with which the values occur ($\sum_{i=1}^{n} p_i = 1$). The generalisation to an enumerable infinity of values x_i and to a continuous variable x is usually not problematic.

Important values for the characterisation of a random variable X are the *expectation value* or the *mean value*

$$\langle X \rangle := \sum_i p_i x_i \tag{1.104}$$

and the *dispersion* or the *mean square deviation*

$$\mathrm{var}(X) = (\Delta X)^2 := \langle X^2 \rangle - \langle X \rangle^2 = \langle (X - \langle X \rangle)^2 \rangle \; , \tag{1.105}$$

which is also called the *variance* var(X). The *standard deviation* $\Delta X = \sqrt{\mathrm{var}(X)}$ indicates how widely the random variable is distributed around its mean value. In quantum mechanics, ΔX is also referred to as the *uncertainty* of X.

1.4 Complementary Topics and Further Reading

- Most textbooks of quantum mechanics contain a summary of the mathematical fundamentals. We mention in particular the following books: [Sak 85], [Ish 95], [Bal 98], [Gri 02], [CDL 05].

- A detailed treatment of Hilbert space with reference to quantum mechanics can be found in [Jor 69].

- The bra space as the vector space of all linear continuous functionals in a vector space V (also called the dual space V^*): [FK 98, Chaps. 2.8 and 4.2].

- A collection of references for Sect. 1.3: [Ish 95], [NC 00].

1.5 Problems for Chapter 1

Prob. 1.1 [for Sect. 1.1]: Prove the relations (1.5), (1.6), (1.7), (1.8), (1.24), (1.25), (1.27), (1.28), (1.30), (1.35), (1.50), (1.61).

Prob. 1.2 [for Sect. 1.1]: Prove that a linear operator which acts on a finite-dimensional complex vector space has at least one eigenvector and one eigenvalue.

Prob. 1.3 [for Sect. 1.1]: Give several examples of a basis set of $\mathcal{H}_3$.

Prob. 1.4 [for Sect. 1.1]: Let $\{|i\rangle, i = 1, \ldots, d\}$ be an ONB. Prove that Parseval's identity

$$\|\varphi\|^2 = \sum_{i=1}^{n} |\langle\varphi|i\rangle|^2 \tag{1.106}$$

holds for all vectors $|\varphi\rangle \in \mathcal{H}_2$.

Prob. 1.5 [for Sect. 1.1]: Show that the matrix which corresponds to the operator product AB is equal to the product of the matrices of A and B.

Prob. 1.6 [for Sect. 1.1]: Show that every linear operator C can be written in the form

$$C = R + iI \tag{1.107}$$

with Hermitian operators R and I. Consider the analogy: linear operator $\leftrightarrow$ complex number; Hermitian operator $\leftrightarrow$ real number.

Prob. 1.7 [for Sect. 1.1]: Show that the determinant of a unitary matrix is ± 1.

Prob. 1.8 [for Sect. 1.1]: Show that for two unitary $n \times n$ matrices U_1 and U_2, the matrix $\begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix}$ is also unitary.

Prob. 1.9 [for Sect. 1.1]: Does the projection operator $P = |u\rangle\langle u|$ have an inverse?

Prob. 1.10 [for Sect. 1.1]:

a) The operator A is known to be diagonalisable. How can its spectral representation be found?

b) Are the Pauli operators $\sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0|, \sigma_y = -i|0\rangle\langle 1| + i|1\rangle\langle 0|$, and $\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|$ diagonalisable? Find their spectral representation.

Prob. 1.11 [for Sect. 1.1]: Give an example of a normal operator which is neither Hermitian nor unitary.

Prob. 1.12 [for Sect. 1.2]: Confirm that the relation

$$\mathrm{tr}[C(\mathcal{A}B)] = \mathrm{tr}[(\mathcal{A}^{-1}C)B] \tag{1.108}$$

holds for the superoperator A defined in Eq. (1.85).

Prob. 1.13 [for Sect. 1.2]: Let H be a Hermitian operator which obeys the eigenvalue equation

$$H|e_i\rangle = E_i|e_i\rangle \ . \tag{1.109}$$

Find the eigenvectors and the eigenvalues of the Liouville operator $\mathcal{L}$ from Eq. (1.87).

Prob. 1.14 [for Sect. 1.2]: Show that the Liouville operator from Eq. (1.87) has the matrix representation

$$\mathcal{L}_{ij,i'j'} = \frac{1}{\hbar}(H_{ii'}\delta_{j'j} - \delta_{ii'}H_{j'j}) \ . \tag{1.110}$$

Prob. 1.15 [for Sect. 1.2]: Prove the following relation by referring to the definition of the Liouville operator $\mathcal{L}$:

$$e^{c\mathcal{L}}A = e^{\frac{c}{\hbar}H}Ae^{-\frac{c}{\hbar}H} \ . \tag{1.111}$$

Prob. 1.16 [for Sect. 1.2]: Describe some situations which can be used to give an intuitive understanding of conditional probability, of the total-probability theorem, or of Bayes' theorem.

© 2007 WILEY-VCH Verlag GmbH & Co.

2 Basic Concepts of Quantum Theory

2.1 First Version of the Postulates (pure states of isolated quantum systems)

2.1.1 Introduction: the Scenario of Quantum Mechanics

If one wishes to apply quantum theory to less well-understood situations, which for example occur in connection with composite systems, then it is expedient first to recall the basic mathematical and physical structures of quantum theory. The present Chapter 2 does just that.

> *We make the* **general physical assumption** *that we investigate only those processes which do not require a relativistic description and which can be formulated in a finite-dimensional Hilbert space.*

The two-slit experiment The characteristic features of quantum physics become clear when they are compared with those of classical physics. To this end, we can consider two analogous physical situations. In the one case, the situation can be described satisfactorily within the framework of classical physics; in the other, a quantum-mechanical description is necessary. The two-slit experiment is a concrete example which is often used in such a comparison; many elements of quantum theory can be illustrated by it. We shall therefore treat this experiment in some detail. The results are intended to prepare the reader for the introduction of the postulates in Sect. 2.1.3 and for the concepts treated in Chap. 4. Entanglement will be discussed later in terms of other experiments, which will thus extend the scenario of quantum theory.

We first describe the experimental situation of a double slit with slits denoted by 1 and 2. In front of the slits, i. e. at the left in Fig. 2.1a, is an apparatus which shoots small balls and which can be controlled by the toss of a coin. Depending on whether the coin shows "heads" or "tails", the apparatus shoots a ball towards slit 1 or slit 2. We assume a uniform scatter of the paths of the balls within the opening of each slit. An observation screen is set up behind the double slits, on which the impacts of the balls can be registered. We discuss the case that only one of the two slits is open (the other is covered), and the case that both are open. For all three cases (only 1 open, only 2 open, 1 and 2 open), we plot the relative frequency of impacts on the screen as a function of position. The greater the number of shots, the more clearly can we discern the relative frequency distributions. When only one slit is open, the distribution appears like that shown in Fig. 2.1a. In the limiting case of a large number of shots, it reflects the probability $P(x)$ of impacts on the screen. If slit 1 is covered, we find the corresponding distribution behind slit 2. It confirms our daily experience that when both slits are open, the

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

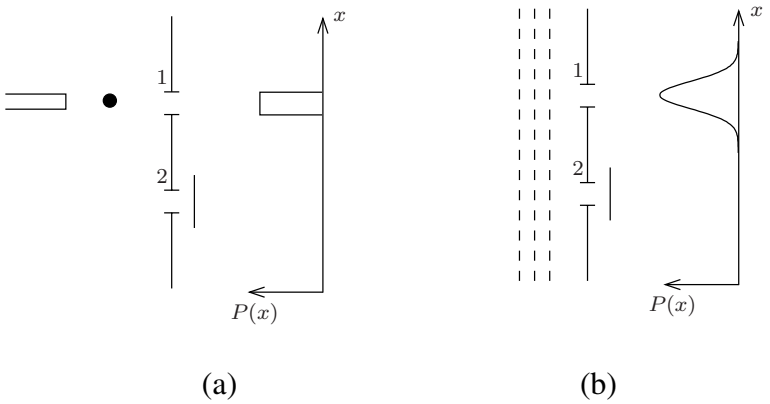


Figure 2.1: A screen behind a single slit: the probability distribution $P(x)$ on the screen for classical objects (a) and for quantum objects (b).

impact probabilities for the individual slits, multiplied by a factor $\frac{1}{2}$, add to give the overall probability. If necessary, the distribution curves can be normalised.

In the analogous quantum-mechanical experiment, the "ball shooter" is replaced by an apparatus containing an atomic-beam source[1]. One can set up suitable apparatus and correspondingly prepared screens (detectors), so that the following condition is met: if the screen is set up without the slits, then impacts are registered with a random distribution over the whole screen. If one then waits long enough, a homogeneous distribution of the impacts is obtained. Since the impacts are individual events, separated in time, we will assume that an individual quantum object (quantum system), which we have already called an atom, has left the source and landed on the screen. We can make no statements about an atom between the source and the screen. We now place a suitably-dimensioned double slit between the source and the screen and close, for example, slit 2. Then, in the limit of a large number of events, we observe the relative frequency distribution (and thus the probability distribution $P(x)$) shown in Fig. 2.1b. Its maximum lies directly behind the opening of the slit. If slit 1 is closed, we find the correspondingly shifted curve behind the open slit 2. However, if we use atoms and open both slits, we observe the distribution of relative frequencies of impacts shown in Fig. 2.2, whose maximum lies directly behind the bridge between the two slits. Again, the limiting case for a large number of events is shown in the figure. As with the balls, on repeating the experiment, the ordering of the positions of the individual impacts varies *in a completely random way* (compare Fig. 2.3). Only in the limit of a large number of events do we always find the same frequency distribution in a *deterministic manner*.

As the essential result, we conclude: for atoms, we obtain the probability distribution from the double slit – in contrast to the case of balls – not by adding the probability distributions from the individual slits; instead, an *interference pattern* has appeared, like that familiar from optics, which cannot be explained by attributing paths to the individual atoms as we could

[1]The two-slit experiment has been successfully carried out with electrons, atoms, van der Waals clusters, Fullerenes, and biomolecules (cf. Sects. 2.6 and 15.6).

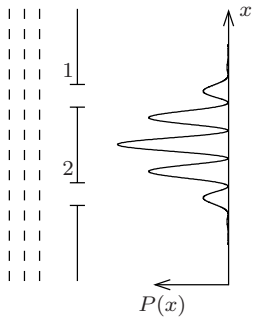


Figure 2.2: A screen behind a double slit: the impact probability $P(x)$ for quantum objects.

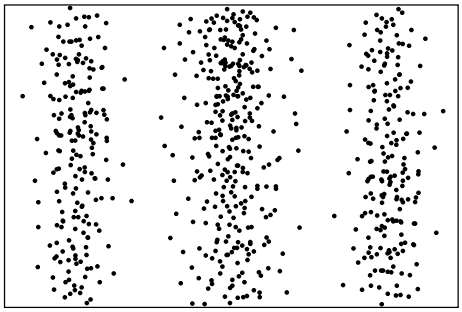


Figure 2.3: It is a completely random event which determines where an individual quantum object strikes the screen. The fringe pattern from a large number of impacts is, in contrast, precisely determined.

for the balls. Owing to the startling analogy to optical diffraction, we can presume that the *mathematical* computation of the probability distribution from the double slit will reflect the phenomenon of interference by superposition in a similar way as in the optical situation.

Either-or *versus* neither-nor For later reference to terms which we shall often use, we need to characterise the experiments with balls or with atoms somewhat more precisely. We summarise the apparatus for shooting the balls, controlled by tossed coins, and the slits in front of it under the term *preparation procedure*. What is prepared is the corresponding *state* of the balls. If only slit 1 (or 2) is open, we will call the state S_1 (or S_2). For each state, the probability distribution of impacts on the screen is well determined. On opening both slits, we prepared a new state, which we will call $S_m(1,2)$. The associated probability distribution is found by addition of the distributions from the *classical states* S_1 and S_2, which are each multiplied with the relative probability $\frac{1}{2}$. Due to the random-number control of the shooting apparatus by tossed coins, balls in state S_1 or in state S_2 are prepared with the same relative probability $\frac{1}{2}$. By mixing the states S_1 and S_2 with this relative weight, we obtain the state $S_m(1,2)$. We call states prepared in this way *statistical mixtures* or also *blends*. A single ball is in either S_1 or S_2. Statistical mixtures are in this sense *either-or states*. The statistical

mixture contains a statistical element owing to the coin tossing. The path of an individual ball is, however, completely determined.

For atoms, we can prepare a *quantum state* $\hat{S}_1$ or $\hat{S}_2$ by closing one of the slits. With many atoms in these states, we obtain again uniquely either the probability distribution of Fig. 2.1b or the correspondingly-shifted distribution from the other slit. If we cover one of the slits with a probability $\frac{1}{2}$, the quantum state $\hat{S}_m(1,2)$ results. The associated probability distribution is found as in classical physics by addition of the probability distributions from $\hat{S}_1$ and $\hat{S}_2$, weighted by the factor $\frac{1}{2}$. Again, we have only mixed the states. This result in quantum theory, as in classical physics, is termed a *statistical mixture* or a *blend*. Since always only one of the slits was open, we can say in this case that the individual atoms must have passed either through slit 1 or through slit 2. We have a quantum-mechanical either-or state. Thus far, there is a perfect analogy to the states of the (classical) balls.

For atoms, however, there is an additional state, $\hat{S}_p(1,2)$, which leads to a probability distribution that cannot occur for the balls. It is prepared when both slits are opened (cf. Fig. 2.2). It is important to note that its probability distribution cannot be prepared by mixing. We call an unmixed state a *pure state*. $\hat{S}_p(1,2)$ is an example. In contrast to balls in the state $S_m(1,2)$, an individual atom behind the double slit is neither in the state $\hat{S}_1$ nor in the state $\hat{S}_2$. The pure state $\hat{S}_p(1,2)$ is a *neither-nor state*. We can therefore not say of an atom that it passed through the one or the other slit. This classical concept, derived from experience with classical objects like balls, fails for atoms. We mention also that the unmixed states $\hat{S}_1$ and $\hat{S}_2$ are likewise pure states according to our definition.

Selective and non-selective measurements By means of measurements, we wish to determine for the mixed states S_m(1,2) and $\hat{S}_m(1,2)$ of balls or of atoms, and for the pure state $\hat{S}_p(1,2)$ of the atoms, precisely which slit an individual ball or an individual atom has passed through. The possible results of such measurements are thus "behind the first slit" and "behind the second slit". For the measurement, we use a light beam which can be scattered by the balls or the atoms directly behind the slits. We observe that a light flash, corresponding to a scattering event, is always seen behind only one of the two slits. We have thus carried out the measurement as desired. Thereafter, the ball or the atom strikes the screen and is registered as before. (In point of fact, the experimental setup is more complicated than we have described here. References to literature on this topic can be found in Sect. 8.8.)

Which frequency distributions of the impacts on the screen do we observe, if we again wait until a large number of events has occurred? The answer depends not only on the state which we measure, but also on the method which we use for evaluating the data from the measurements (see Table 2.1). A possible procedure consists of making a conditional measurement and registering only those impacts which are associated for example with flashes of scattered light behind slit 1. We call this a *selective measurement* (or a measurement with post selection). Such a measurement consists of many events with subsequent selection depending on the result of the slit determination. We observe from the resulting impact-frequency distribution the following: the classical statistical mixture in the state $S_m(1,2)$ passes over to the state S_1. Analogously, the quantum-mechanical statistical mixture $\hat{S}_m(1,2)$ passes over to the

	Classical Physics		*Quantum Physics*	
Selective Measurement	Statistical Mixture	→ into one of the pure states contained in the mixture	Statistical Mixture	→ into a pure state. $(*)$
			Pure State	→ into an (in general different) pure state
Non-selective Measurement	Statistical Mixture	→ into the same statistical mixture	Statistical Mixture	→ into a possibly different statistical mixture
			Pure State	→ into a statistical mixture

Table 2.1: The effects of measurements in classical physics and in quantum physics. The arrows → mean: "..." is transformed by the measurement into "...". For a statistical mixture, we also use the term blend. As we shall later see, (*) holds only when no degeneracy is present.

state $\hat{S}_1$. Both results are not surprising. We have simply separated the mixed states again by selective measurement.

With atoms, we can in addition prepare the pure state $\hat{S}_p(1,2)$. A selective measurement belonging to flashes of scattered light from behind slit 1 transforms this state according to the observed impact-frequency distribution into the state $\hat{S}_1$. In the case of the double slit experiment, the selective measurement thus transforms a pure state into a different pure state. The measurement acts on the states and changes them. We can also interpret a selective measurement as a *re-preparation* of the states. In special cases, there is no change in the state: if the pure state $\hat{S}_1$ is already present (slit 2 is closed), then flashes are observed only from behind slit 1. Furthermore, the frequency distribution then observed from the atoms that produced flashes demonstrates that the state $\hat{S}_1$ was not changed. In this case, the measurement just verifies the state preparation.

An alternative procedure for data evaluation consists of not making a selective measurement, i. e. we register the scattered light from the atoms or the balls, but collect all the impacts on the screen into a single image, without taking into account where the scattered-light flashes originate: we thus carry out a *non-selective measurement*. From the resulting probability distribution, we can determine that for both balls and atoms, a statistical mixture is again transformed into a statistical mixture. This is plausible; since we made no selection, the states remain mixed. For atoms, we can in addition prepare the pure state $\hat{S}_p(1,2)$, which has no corresponding state for the balls, When we measure non-selectively from this state, we obtain the impact-frequency distribution belonging to the statistical mixture $\hat{S}_m(1,2)$. We have mixed the states $\hat{S}_1$ and $\hat{S}_2$ belonging to the two possible results of the slit determination. In quantum physics, a non-selective measurement transforms a pure state into a statistical mixture.

Finally, we want to make a last remark concerning measurements with atoms. We already described how the state $\hat{S}_1$ can remain unchanged by a measurement. A second measurement immediately following the first is therefore performed again on the state $\hat{S}_1$, and we register

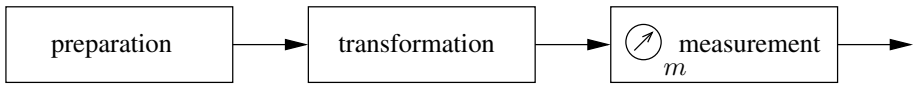


Figure 2.4: The actions of experimental apparatus on a physical system.

again a flash behind slit 1. This is the reason why one can attribute to atoms which caused a flash for example behind slit 1 the *property* "behind slit 1". A measurement which repeats the question "behind which slit?" leads for these atoms again to the same answer, "behind slit 1".

The typical experimental situation We can intervene in other ways than by a measurement between the double slits and the screen. If the balls or the atoms are electrically charged, we can for example apply an electric field, which will distort the image of the impact probabilities. The weaker this field, the less distortion. The electric field produces a change in the state of the objects. This is only one example; there are other ways to intervene in the experiment. We collect them all under the concept of the effects of *transformation apparatus* which transform the state. The undisturbed free evolution of the state is included as a special case.

In summary, we can say that physics, including quantum physics, deals with three types of macroscopic apparatus: preparation apparatus, transformation apparatus, and measurement (or detection) apparatus. In an experimental setup in a laboratory, a particular sort of each of these types of apparatus is used. The experiment can be decomposed into three independent sequential phases: preparation, deterministic transformation, and non-deterministic (probabilistic) measurement (see Fig. 2.4). The measurement can be evaluated selectively or non-selectively. The corresponding state following the measurement is in general different from that prior to the measurement.[2]

Of course, quantum mechanics can also be applied to the description of processes on distant stars or in earlier stages of the universe. The concepts of preparation and measurement must then be suitably modified.

The domain of application of quantum mechanics The goal of a physical theory is to predict the results of a measurement which a measurement apparatus will yield when the effects of the preparation and the transformation apparatus are known, and to justify these predictions within the framework of the theory. We saw already in the example of the two-slit experiment – in particular for the state $\hat{S}_p(1,2)$, which we characterised as the result of an "interference process" – that in certain cases there is no justification within the framework

[2] In the case of the two-slit experiment, we dealt only with the preparation and the measurement. The fact that the result shown in Fig. 2.2 has the form of an interference pattern is not to be misunderstood in terms that the individual quantum object takes on the form of a wave or behaves like a wave. In the Heisenberg representation of quantum mechanics, there is no recognisable wave propagation even in the mathematical formulation. Similarly, the fact that a quantum object produces a point on a screen does not imply that it behaves as a particle. The measurement apparatus is simply constructed in such a way that it displays point impacts. If only one slit is open, the quantum object can only have passed "along this path". This again does not imply particle behaviour. The so-called quantum-mechanical *wave-particle problem* (or *wave-particle duality*) is a pseudo-problem. The fact that one can, on the other hand, reasonably speak of individual quantum objects is due to the possibility of carrying out preparation and measurement processes which are resolved on the time axis and can be individually controlled.

of classical physics. These experiments set a limit to the domain of application of classical physics. Their justification lies outside this domain, in that of *quantum physics*. A *quantum effect* is present when it is not possible to give a purely classical explanation of the behaviour of the three types of macroscopic apparatus. Physics attempts to reach objective conclusions. An unambiguous intersubjective communication requires information exchange via classical media (cf. Sect. 6.4.2) and must be based on facts which can be described by the methods of classical physics. This applies to the preparation apparatus, the transformation apparatus, and to the measurement apparatus. They are all *interfaces* between the classical domain and the quantum domain, and in this sense they are not purely classical instruments.

Quantum theory explains quantum effects. Whether and in what way it can also give explanations for the effects of classical physics, i. e. whether classical physics can be derived as a limiting case from quantum theory, is an open question. It is the subject of current research, and we will return to this question in Chap. 15. The reverse conclusion is also conceivable. The domain of application of quantum theory could be empty. Can quantum effects perhaps be described by classical physics? We will take up this question again in Chap. 10, where we deal with "hidden variables".

2.1.2 Quantum States

In the following section, we will reduce the theoretical justification of the phenomena in the quantum domain to a few basic assumptions. To this end, we generalise the experience which we gained from the atomic-interference experiment. We shall not, however, attempt to achieve the mathematical and conceptual precision of an axiomatic quantum theory. We refer the reader to the literature for such an axiomatic treatment. Here, we begin with a more precise rendering of the concept of the quantum state, of which we have already made use, and then introduce the postulates which can explain the quantum phenomena that occur in the case of pure states. In the later chapters of this book, we will reformulate these postulates in a more and more general way. In the process, it will become clear that the basic structure which is determined in the end by the scheme shown in Fig. 2.4 remains valid (cf. Fig. 14.4).

Quantum systems The apparatus in figure 2.4 act sequentially in the order from left to right. The arrows indicate the transfer of the quantum system (quantum object) from one apparatus to the next. The actions of the apparatus define the system. The term *quantum system* denotes a system which has passed through a preparation process and on which measurements can be performed. The justification of the results of the measurements must then fall in the domain of quantum theory. In the system discussed above, an individual atom is such a quantum system. The spin orientation of an atom or the polarisation of a photon can also be prepared and detected. *In the* standard interpretation *of quantum mechanics, which we employ in this book, physical* reality *is attributed to individual quantum systems* (cf. Footnote 2). Their existence is postulated, just as we have already done, based on the individual impacts on the screen, in the case of the atomic interferometer. The numerous experiments with single ions in Paul traps and individual Rydberg atoms in microwave resonators render this postulate plausible. We shall return to the problem of reality in Sect. 2.5. As we shall see, quantum systems

can themselves be composed of subsystems. In this case, entangled states play a central role (cf. Chap. 7).

Quantum states and measurements Measurements on quantum systems which were prepared in an identical manner can lead to different results. For example, in our two-slit experiment, the atoms can be detected at different positions on the screen (see Fig. 2.3). A particular preparation determines only the probabilities of various measurement results. In order to determine the probability distribution experimentally, measurements must be carried out on a very large number of identically-prepared systems. *The state of a quantum system is associated with a particular preparation procedure. The term* quantum state *refers to that particular mathematical (!) object which permits us to calculate unambiguously the probabilities of the results of all possible measurements on systems which have passed through the associated preparation procedure.* We thus do not expect that a quantum state introduced in this manner has a counterpart in reality which is "carried" by an individual quantum system. There is in general no property of a quantum system which corresponds to its state vector. If one wishes for clarity to specify something corresponding to a "quantum state", then it should be a preparation apparatus (or more precisely, a preparation procedure). Numerous misunderstandings can be avoided by keeping this interpretation of the term "quantum state" in mind.

In contrast, in the case of classical objects, the mathematical expression for the state refers to the particular object itself and thus has a direct correspondent in reality. The state "cold and transparent" can be physically realised by an individual ice cube. The concept of "state" in quantum physics is very different from that in classical physics, at least for pure states. As we shall see in Chap. 4, in the case of statistical mixtures, there are to some extent similarities.

Different preparation procedures can lead to the same state. These procedures form in this sense an equivalence class of state preparations. The mathematical description of the state and the computation of the probabilities will be postulated in the following. By referring to the equivalence classes, the individual structures of particular preparation and measurement apparatus of the same type will be eliminated.

It is important in terms of this background to clarify a *manner of speaking* which has become common and which we will also use for simplicity: one frequently says that a single quantum system is to be *found in* a certain state, or *has* this state. *This means that it has passed through a preparation procedure belonging to a particular equivalence class of preparations.* Only in this sense can one attribute a certain state to an individual system. One also says that a single quantum system is *transformed* from one particular state into another state as a result of an external influence, e. g. due to a measurement. This means that the quantum system has passed through a preparation process which includes the influence of the measurement. The resulting state again makes a statement only about the probability distribution of the results of many measurements on quantum systems which have been prepared in this manner.

We thus remain very reticent in terms of statements about single quantum systems in this first step towards the formulation of the standard interpretation of quantum mechanics. We will draw more extensive conclusions about the existence of properties such as energy, position, etc. only after formulating the postulates.

Isolated quantum systems As in the classical mechanics of a "free point mass", the concept of a *free system* is fundamental to the structuring of the quantum theory. This concept is an idealisation, which in fact can be attained only approximately as a limiting case. It is based on the idea that in certain situations, quantum systems can be so completely disconnected from the rest of the world that all the possible processes in that rest leave the state of the system unaffected. The system cannot be modified prior to the measurement.

Free quantum systems are not of interest for applications. We thus allow the state of the system between the time of its preparation and the time of measurement to evolve in a different manner than a free state. As in classical mechanics, in quantum mechanics a cause is attributed not to free-system behaviour itself, but only to deviations from free behaviour; here, these are represented by the transformation apparatus. It describes external constant or time-dependent influences during the evolution of the system between preparation and measurement, as they are caused e. g. by classical electromagnetic or gravitational fields. We first consider the special case that the external systems are not themselves affected. The quantum system is thus supposed to be isolated with respect to reactions "to the outside world". This can be only approximately achieved in practice. Furthermore, the mediators of the external influences from the transformation apparatus (for example the electric field) can be effectively described by classical physics, i. e. they have no quantum-mechanical degrees of freedom of their own. They are represented by operators which form parts of the Hamiltonian (compare Eq. (2.10)). When these two approximations are fulfilled, one usually refers to the quantum systems as *isolated quantum systems* (or *closed systems*). This term is not particularly well chosen. Non-isolated systems are called *open systems*. We will discuss them after introducing composite quantum systems in Chap. 7.

Of course, isolated systems are open to measurements. A reaction onto the measurement apparatus then takes place. In contrast to the classical fields, the measurement apparatus has also quantum-mechanical degrees of freedom. We will return to this point in Chap. 15.

Pure states The preparation apparatus can itself be composed of other preparation devices, which act with well-determined frequencies and carry out different preparations of the quantum systems. In this case also, the unambiguous prediction of the probabilities of all measurement results is possible. The overall preparation apparatus prepares a state which in view of the many contributing preparation procedures is called a *statistical mixture* (or *blend*). The state $\hat{S_m}(1,2)$ described in Sect. 2.1.1 is an example. We will investigate states prepared in this particular way in detail later. They are special mixtures. The adjective "statistical" emphasizes this point. For the initial version of the postulates, we exclude mixtures. We restrict ourselves to states which can not in any sense be produced by a mixing procedure or, with respect to the associated probability statements, be simulated by a mixture. As in Sect. 2.1.1, we refer to these states as *pure states*. Examples are the states $\hat{S}_1$, $\hat{S}_2$ and $\hat{S}_p(1,2)$ described in Sect. 2.1.1. Along with the isolation of the system, the purity of the states is the second major idealisation which we initially require. As in classical mechanics, which is based upon a postulate for free point masses (inertial systems), we will then advance step by step towards a description of realistic physical situations.

2.1.3 Postulates for Pure States of Isolated Quantum Systems

We now have all the concepts at our disposal which we require to formulate a *first version* of the postulates. We will generalise all three postulates in later chapters.

Postulate 1 (pure state) *An isolated quantum system which is in a pure state is described by its* state vector $|\psi\rangle$. *This is a normalised vector in a Hilbert space $\mathcal{H}_d$ which is associated with the quantum system.*

We first simplify our basic experimental situation and pass directly to the measurements. For this purpose, we suppose that the transformation apparatus has been removed from the experiment or else we include it as a part of the preparation apparatus. The measuring device will not at this point be treated in the most general way. Instead, we limit ourselves to *projective measurements*. These are also called *von Neumann measurements*. This is a fundamental type of measurement which will repeatedly play a central role in our later generalisations. This type of measurement is characterised by its action on the state vector.

Postulate 2 (projective measurements, non-deterministic dynamic evolution)

a) *A projective measurement of a physical quantity (e. g. of the energy, angular momentum, etc.) carried out on a quantum system is described by an Hermitian operator which can be time dependent and acts on the vectors of $\mathcal{H}_d$. We speak of a measurement of the* observable *A and denote the operator with the same symbol A.*

b) *The possible measured values which can occur as a result of a measurement of the observable A are the eigenvalues a_n of the associated operator A. For simplicity, we assume that its spectrum is discrete:*

$$A|u_n^i\rangle = a_n|u_n^i\rangle, \quad i = 1, \ldots, g_n \; . \tag{2.1}$$

The eigenvectors $|u_n^i\rangle$ form an orthonormal basis or can, in the case of degeneracy, be correspondingly chosen. The g_n give the degree of degeneracy *of the eigenvalues a_n.*

c) *When a selective measurement of the observables A of a system with a normalised state vector $|\psi\rangle$ leads to the result a_n, then the non-normalised state vector $|\tilde{\psi}'_n\rangle$ immediately following the measurement is given by the projection of $|\psi\rangle$*

$$|\psi\rangle \to |\tilde{\psi}'_n\rangle = P_n|\psi\rangle \tag{2.2}$$

with the projection operator

$$P_n := \sum_{i=1}^{g_n} |u_n^i\rangle\langle u_n^i| \; , \tag{2.3}$$

which projects onto the space of the eigenvectors corresponding to a_n. Through normalisation of $|\tilde{\psi}'_n\rangle$, the state vector $|\psi'_n\rangle$ after the measurement is obtained.

d) *We denote by $N(a_n)$ the frequency with which a measured value a_n is obtained when the measurement is carried out on N identically prepared systems in the state $|\psi\rangle$. The relative frequencies $\frac{N(a_n)}{N}$ for all these ensembles approach the probability $p(a_n)$ as a limiting value in the limit $N \to \infty$:*

$$\frac{N(a_n)}{N} \xrightarrow{N\to\infty} p(a_n) \,. \tag{2.4}$$

e) *The probability $p(a_n)$ of obtaining a particular measured value a_n at a certain time is equal to the expectation value of the projection operator P_n computed with the state $|\psi\rangle$ prior to the measurement. Equivalently, it is equal to the square of the norm of the non-normalised state vector $|\tilde{\psi}'_n\rangle$ after the measurement:*

$$p(a_n) = \langle\psi|P_n|\psi\rangle = \|\tilde{\psi}'_n\|^2 \,. \tag{2.5}$$

The last equation results from Eqs. (1.4) and (2.2). Since A is an Hermitian operator, $\sum_n P_n = \mathbb{1}$ and therefore, as is to be expected of a total probability,

$$\sum_n p(a_n) = \langle\psi|\psi\rangle = 1 \,. \tag{2.6}$$

For the expectation value or mean value, we obtain using the dyadic decomposition of A:

$$\sum_n p(a_n) a_n = \langle\psi|A|\psi\rangle \,. \tag{2.7}$$

We return to the topic of quantum measurements in more detail in Chap. 13.

Finally, we describe the effect of the intervention of the transformation apparatus on isolated systems:

Postulate 3 (deterministic dynamic evolution between preparation and measurement)

a) *For isolated systems, the probability distribution $p(a_n)$ evolves in a deterministic and reversible manner between the preparation and the measurement. Its time development between two times t_0 and t_1 is described by a unitary* time-development operator *$U(t_1, t_0)$:*

$$U^\dagger(t_1, t_0) = U^{-1}(t_1, t_0) \,. \tag{2.8}$$

This operator fulfills the conditions $U(t_0, t_0) = \mathbb{1}$ and

$$U(t_2, t_1)U(t_1, t_0) = U(t_2, t_0) \tag{2.9}$$

for arbitrary times t_0, t_1, t_2.

b) *The dynamic equation for $U(t, t_0)$ is*

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} U(t, t_0) = H(t)U(t, t_0) \,. \tag{2.10}$$

It is formulated with an Hermitian operator H (the Hamiltonian). $\hbar = 1,0546 \times 10^{-34}\text{Joule} \cdot \text{sec}$ is Planck's constant. It is postulated that $H(t)$ is the observable which belongs to the total energy of the system. It can be explicitly time dependent in a time-dependent problem. No other time dependencies are contained in the operator $H(t)$ in Eq. (2.10).

c) *The* Schrödinger representation *is one of the many possible formulations of this time development. In this representation, the dynamic evolution of the state is given by the state vector alone, according to*

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle \ . \tag{2.11}$$

Observables can be only explicitly time dependent. At each time t, there is a corresponding probability distribution $p(a_n, t)$ for the results of a measurement of A given by Eq. (2.5). From Eqs. (2.10) and (2.11), the Schrödinger equation *follows*:

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H(t)|\psi(t)\rangle \ . \tag{2.12}$$

Other representations, such as the Heisenberg and the interaction representations, can be found by making use of unitary equivalence. This guarantees that all statements about measured values and the probability of their occurrence are the same in all representations at each time. *In general, we shall make use of the Schrödinger representation.*

Why a unitary time evolution? As a result of the unitarity of the time-development operator, the state vector $|\psi\rangle$ remains normalised and the overall probability of observing one of the measured values is equal to one: $\sum_n p(a_n, t) = 1$. If one conversely requires the conservation of the overall probability during the dynamic evolution with a (not necessarily unitary) time-development operator $T(t_1, t_0)$, then at the time t_1, again,

$$\sum_n p(a_n, t_1) = \langle T(t_1, t_0)\, \psi | T(t_1, t_0)\, \psi \rangle = 1 \tag{2.13}$$

must hold for all the states $|\psi\rangle$. As we have shown in Sect. 1.1.5, this conservation of the norm implies the unitarity of T. One could reformulate Postulate 3 and place the physically plausible requirement of conservation of the total probability in the foreground. Then the deterministic evolution can be effected only by a unitary time-development operator.

Physical properties The postulates allow us to expand upon our previous interpretation. To what extent can quantum systems be ascribed with certain physical properties, as is usual in the standard interpretation? When the state is an eigenvector $|u_n\rangle$ with the eigenvalue a_n of the observable operator A, then a measurement of A by a suitably-constructed measurement device leads with certainty to the result a_n. If we repeat the measurement, it will with certainty always yield a_n. This is the characteristic property of projective measurements, due to $P_n P_n = P_n$.

It is therefore reasonable to say that the system prepared in the state $|u_n\rangle$ possesses the *physical property* A which has the value a_n. It is assumed to be real. If A is for example

the energy operator, then the system *has* the energy a_n; the corresponding conclusion holds for the angular momentum operator, etc. If for a general state $|\psi\rangle \neq |u_n\rangle$ the measured value a_n is obtained from a measurement of A, it is in general not safe to say that the system *previously* had the property a_n. Only through the interaction of the quantum system with the measurement apparatus is the system transformed into the state $|u_n\rangle$ so that the measurement apparatus always yields the value a_n on repeated measurement of A. The concept of property and of measurement thus differ in essential ways from the like-named concepts in classical physics. This will play a role in particular in Chap. 10. Later, in connection with generalised measurements (Chaps. 13 and 14), we will go beyond the concept of an observable represented by an Hermitian operator. In Sect. 2.4, we return to the interpretation of quantum theory as summarised above.

2.1.4 Comments on the Postulates

- The dimension of the Hilbert space of a quantum system is physically characterised as the maximum number of states which can reliably be distinguished from one another in a single measurement. This becomes clear if one takes as observable an Hermitian operator, none of whose eigenvalues is degenerate.

- Along with the dimensionality of the Hilbert space, there are additional characteristics of quantum systems which are independent of the observer and are not subject to probability statements. Among them are the *classical variables* mass, charge and the magnitude of the spin of a quantum system. Although these quantities can be determined experimentally, they appear in non-relativistic quantum mechanics only as parameters. They are also called *intrinsic properties*.

- *Time* is a classical variable in quantum theory and not an observable. Clocks are a part of the classical apparatus in Fig. 2.4. A measurement of time is a classical measurement. The passage of time is registered in principle by the motions of free classical point masses. Other procedures (e. g. atomic clocks) are in the end derived from such processes.

- According to postulate 3, the probability statements $p(a_n, t)$ evolve between preparation and measurement in a *deterministic* and *reversible* manner. The corresponding quantum events (the occurrence of measurement results) are in contrast not predictable in individual cases. Only their probability distribution $p(a_n)$ is known at every given time. The evolution during the measurement is *non-deterministic* and *irreversible*. The transition from Eq. (2.2) is often called a *state reduction*. With the new state $|\psi'_n\rangle$, the probability predictions are different. This however does not imply that a physical process takes place in space and time which could reasonably be denoted as a *collapse of the wavefunction*.

- The deterministic process represents an observable change with time in the probability distribution. In the Schrödinger representation, this is described by a time-dependent state vector $|\psi(t)\rangle$. If the evolution described by a unitary operator were interrupted at a time t and a measurement then carried out, then the probabilities for the measurement

results could be computed with the mathematical ancillary quantity $|\psi(t)\rangle$ from postulate 2. In the Heisenberg representation, in contrast, the state vectors are independent of time and the observable operators are time dependent. This again makes it clear that the state vector does not represent a description of the system, such as is familiar from the concept of 'state' in classical physics.

- The fact that the time-development operator $U(t, t_0)$ depends on the time as a continuous variable should not mislead one into thinking that a quantum system evolves continuously with time, e. g. in terms of its properties. The transition of an atom from the source to the detector in the two-slit experiment is an overall process which is not further divisible. *Between preparation and measurement, there are no further events.* This is a characteristically discrete element of quantum mechanics. *The time-development operator describes the time-dependent changes in the probabilities of possible facts, but not the changes in the facts themselves.* In this connection, greatly simplified manners of speaking are often used, e. g. "an atom passes through the slit". If conceptual problems or paradoxes are generated by this language, one must return to a more precise formulation and its operational significance.

- It is the mathematical coaction of the state vector and the observables (compare Eq. (2.5), which is valid in all representations) that maps the interaction of the quantum system with the measurement apparatus in the laboratory. In this process, not only does the measurement device enter a new state, as in classical physics, but so does the quantum system also. The system as a rule must be described by a new state vector after the measurement (cf. Eq. (2.2)).

- It is permitted that the measurement destroys the quantum system. Then, part c) of postulate 2 can be dispensed with.

- It is assumed that not only every pure state can be represented by a state vector, but also that every state vector represents a physically-possible pure state. The associated preparation procedure can in principle be experimentally implemented.

- We introduced the pure state as a state which cannot be produced by mixing. This negative characterisation is useful to only a limited extent in practical applications. However, in postulate 2.c, we described a procedure for distinguishing a pure state which can more readily be operationalised and which we can fall back on: *a pure state is found as the result of a measurement when the measured values are not degenerate.* In the case that degeneracy is present, a complete set of pairwise-commuting observables must be measured. The set of associated measured values characterises the resulting state vector unambiguously.

- The postulates presume that in the quantum domain there are *two completely different types of dynamical processes*: the irreversible measurement process with the corresponding probability statements (postulate 2), and the reversible unitary state evolution between the preparation and the measurement (postulate 3). The desire for unification suggests that we could extend the quantum system to include a purely quantum-mechanically describable measurement device, thus obtaining a larger closed quantum

system. One could then attempt to describe the common evolution in the sense of postulate 3. Postulate 2 would become superfluous. Later, we will discuss approaches of this type (cf. Chap. 15).

If this research programme is to be successful and thus yields all of our results, then we would have described here only a *pragmatic approach* in which one for brevity employs classical concepts and two different dynamics as though they were fundamental. For this reason, we can initially retain the conclusion that in postulates 2 and 3, two different dynamics are introduced: the *measurement dynamics* and the *transformation dynamics*, which we will refer to also as *unitary dynamics*.

- We will assume that for every Hermitian operator, a suitable measurement device can be constructed. In fact, in most cases it is by no means a trivial task to design such an experimental implementation.

- We have seen that the concepts

  State
  Measurement
  Property

  are notably different in classical physics and in quantum physics. Many apparent paradoxes and conceptual difficulties arise from the use of the same terms, suggesting too many similarities. For quantum-physical systems, one should rather speak of the quantum state, quantum measurements, and quantum properties. If this is not always adhered to for reasons of economy, the reader should mentally add the virtual prefix "quantum" to the corresponding terms.

- And finally: in order to understand quantum theory, we need to free ourselves from incorrect preconceptions. This includes, to be sure, the false conception of what is meant by "understanding".

2.2 Outlook

In formulating the postulates, we have made use of a number of physical constraints, which we will relax step by step in the coming chapters, until we finally obtain a quite general structure for the quantum theory.

- We have limited ourselves to pure states. The general quantum state is a mixture (Chap. 4).

- Quantum systems can be composed of subsystems, which are then not in general isolated. The quantum theory of such open systems has to be developed (Chaps. 7 and 8). In composite systems, we will meet up with entangled states.

- Projective measurements are a special type of quantum measurements. In Chaps. 13, 14 and 16, we will introduce generalisations of this concept.

- For open quantum systems, dynamic evolutions are possible which are no longer describable in terms of unitary time-development operators. We will formulate them using superoperators (Chap. 14).
- What can be gained by attempting to reduce the measurement dynamics of postulate 2 to the transformation dynamics of postulate 3? We discuss this point in Chapter 15, together with the question of whether classical physics can be derived from quantum theory.

All of these conceptual developments not only deepen our understanding of non-relativistic quantum mechanics, but also lead to new physical effects and form the basis for quantum information theory and for quantum computing. We will provide additional outlooks in Sections 7.3.1 and 8.2.

Further generalisations, which we however cannot discuss here, could also be made: we could include observable operators with continuous eigenvalue spectra, such as position and momentum. If the number of quantum systems is not fixed or well-determined, their description requires the introduction of a Fock space (many-body systems, field quantisation). In both cases, additional new effects can be expected. On generalising to Hilbert spaces of enumerable-infinite dimensionality, in contrast, our results up to now can be directly adopted for those cases with physical relevance.

Before we carry out the generalisation steps which we have listed above, we want to demonstrate in the following the power of projective measurements and, in Sections 2.4 and 2.5, to cast a view from a higher vantage point onto the structure of the theory as described up to now.

2.3 Manipulation of the Evolution of the States by Projective Measurements

Quantum-mechanical measurements intervene in the dynamic evolution of a quantum system and change it. With projective measurements, this intervention is particularly strong. By applying a sequence of projection measurements, we can "freeze" the evolution or impress an arbitrary development onto the state; thus, by measurements, we can 'drive' the system.

2.3.1 The Quantum Zeno Effect

Short-time behaviour We consider the following situation: the state of the system at the time $t = 0$ is an eigenvector $|a\rangle$ of an observable A: $|\psi(t = 0)\rangle = |a\rangle$. A has a discrete spectrum. The unitary development takes place under the influence of the time-independent Hamiltonian H. We set $\hbar = 1$.

$$|\psi(t)\rangle = e^{-iHt}|a\rangle \,. \tag{2.14}$$

After the time t, we measure the observable A. The probability of finding the system after this measurement once again in the initial state $|a\rangle$ is

$$p(t) = |\langle a|e^{-iHt}|a\rangle|^2 \,. \tag{2.15}$$

For short times, we obtain from this

$$p(t) = 1 - (\Delta H)^2 t^2 + \mathcal{O}((\Delta H)^4 t^4) \tag{2.16}$$

with the energy uncertainty ΔH

$$(\Delta H)^2 := \langle a|H^2|a\rangle - \langle a|H|a\rangle^2 =: \tau_z^{-2} \,. \tag{2.17}$$

The time τ_z is called the *Zeno time*. It is the longer, the more similar $|a\rangle$ is to an energy eigenstate. In the limiting case $\Delta H = 0$, we obtain $p(t) = 1$. For $\Delta H \neq 0$, $p(t)$ depends on the square of t at short times, $t \ll \tau_z$.

The quantum Zeno effect [3] We now carry out a series of N measurements of the same observable A in equal time intervals over a total time T

$$\tau := \frac{T}{N} \tag{2.18}$$

with $\tau \ll \tau_z$. The conditional probability $p^{(N)}(T)$ of finding the initial state $|a\rangle$ following each individual measurement in this series each time is, with Eq. (2.16),

$$p^{(N)}(T) = [p(\tau)]^N = [p(T/N)]^N \approx \left(1 - \frac{1}{\tau_z^2}\left(\frac{T}{N}\right)^2\right)^N \,. \tag{2.19}$$

One could also dispense with the restrictive requirement that after each measurement the state $|a\rangle$ should be present. In this case, the probability of finding the system in the state $|a\rangle$ after the measurement at time T is greater than $p^{(N)}(T)$.

The more measurements which are carried out within the time interval $[0, T]$ for a fixed T, i. e. the shorter the time interval τ, the greater is the probability that the system remains in the initial state $|a\rangle$. In the limit $N \to \infty$ or $\tau \to 0$ of a continuous projective measurement, the measurement dynamics completely dominate the unitary evolution and the system is "frozen" into the initial state:

$$p^{(N)}(T) \xrightarrow{N\to\infty} 1 \,. \tag{2.20}$$

This is called the *quantum Zeno effect*. In contrast to Eq. (2.19), the limiting case in Eq. (2.20) is in fact unphysical: the quantum-mechanical measurement process has a certain finite duration.

2.3.2 Driving a State Vector by a Sequence of Projection Measurements

We can not only to a good approximation prevent the evolution of a quantum state by repeated projection measurements, but we can also control its time development. Let the Hilbert space of the system be two-dimensional, with the ONB $\{|\uparrow\rangle, |\downarrow\rangle\}$. The initial state at time $t = 0$ is $|\uparrow\rangle$. For an observable which has an eigenstate that is rotated relative to this initial state,

$$|\alpha\rangle = \cos\alpha|\uparrow\rangle + \sin\alpha|\downarrow\rangle \,, \tag{2.21}$$

[3] Named after the arrow paradox due to Zeno of Elea (ca. 495-430 B.C.).

the probability of finding the system after the measurement in the state $|\alpha\rangle$ is given by

$$p(\alpha) = \cos^2 \alpha \,. \tag{2.22}$$

We again carry out N measurements at time intervals $\tau = \frac{T}{N}$ in the overall time interval $[0, T]$. But in this case, we measure sequentially new observable operators, which have the eigenstates $|\alpha_n\rangle$, with $\alpha_n = n\omega\tau$ and $n = 1, 2, 3, \ldots$. We assume that there is no additional unitary evolution. Then the conditional probability of finding the system in the state $|\alpha_n\rangle$, when it was previously in the state $|\alpha_{n-1}\rangle$, is

$$\tilde{p}(n) = |\langle\alpha_n|\alpha_{n-1}\rangle|^2 = \cos^2 \omega\tau \,. \tag{2.23}$$

The probability of finding the system after each of these measurements in the corresponding eigenstate $|\alpha_1\rangle, |\alpha_2\rangle, |\alpha_3\rangle \ldots$ is given by

$$\tilde{p}^{(N)}(n) = \left(\cos \omega \frac{T}{N}\right)^{2N} \xrightarrow{\text{N large}} 1 - \frac{\omega^2 T^2}{N^2} N \xrightarrow{N\to\infty} 1 \,. \tag{2.24}$$

In the limit $N \to \infty$ for a fixed t or for the time interval $\tau \to 0$, the state of the system always agrees with the state $|\alpha\rangle$ of Eq. (2.21) with $\alpha = \omega t$. The system has been forced to follow a predetermined evolution of state by a sequence of adapted projective measurements. In this case, also, the limit $\tau \to 0$ is unphysical in the strict sense, owing to the finite duration of the measurement process.

2.4 The Structure of Physical Theories*

Thus far, we have spoken only of the standard interpretation. Are there other interpretations? What is meant by the interpretation of a physical theory? We will turn to these questions in the next two sections. Both are not essential for understanding the remaining chapters and can therefore be skipped over. On the other hand, questions connected with problems of interpretation in natural philosophy[4] and in the *philosophy of science*[5] play a large role both in fundamental discussions and in popular-science descriptions. In particular, the question, "What does the quantum theory tell us about reality?" is apparently a source of great fascination for many physicists and non-physicists alike. This justifies some remarks, accompanied by an asterisk, on how such questions are to be dealt with. These can be useful even for the more "practically-oriented" readers, since they can help to avoid a certain confusion in the discussion of quantum-theoretical problems and aid in understanding statements made by quantum mechanics.

*The sections marked with an asterisk * can be skipped over in a first reading.

[4] *Natural philosophy* investigates the concepts that are necessary for understanding the statements which the natural sciences can make about nature. It is an *ontology* with reference to nature, i. e. a doctrine of existence. There are many ontologies.

[5] The *philosophy of science* provides a logical analysis of scientific theories and compares theories with each other. Among the topics treated are the logical structure of elucidations, modes of comparison of theories to experience in the natural sciences, social influences on the formulation of theories, concepts of reality, limits of knowledge, physical theories as guides for actions, reality as a construction, and many more.

2.4.1 Structural Elements of a Physical Theory*

Forerunner theories We close this Chapter 2 on the fundamental concepts of the quantum theory with some structural considerations. Here, we want to consider in particular the concept of reality to which the quantum theory refers. To this end, it is first helpful to clarify just how physical theories are structured. We take a brief excursion into classical physics and consider classical electrodynamics. Typical elementary experiments involve the measurement of forces, heating of wires and similar effects. Measurements of forces, heat and other quantities refer to theories such as mechanics, thermodynamics etc., which were formulated before electrodynamics and independently of it. They are *forerunner theories*. Along with wires and masses, we include force fields, heat etc. within physical reality. These are the elements of reality which were already introduced with reference to the forerunner theories. When we formulate electrodynamics experimentally and theoretically, we presume that the apparatus and measurement devices which are based on the forerunner theories are parts of physical reality.

Structural elements We can already read off several structural elements of a physical theory from the example of electrodynamics. A *physical theory* is a mathematical-deductive scheme which contains the following minimal components:

1. a *mathematical component* MC which consists of mathematical quantities, definitions, equations, transformations, solution procedures etc.;

2. a part from nature, which is called the *basic domain* BD (we will explain this part more precisely below), and which is assumed to exist; and

3. *mapping principles*, which are also called *correspondence rules* CR, which relate the basic domain BD and parts of the mathematical component MC to each other.

Only when these mappings are added to the mathematical relations does a physical theory come into being. Typically, the mathematical quantities are also characterised by dimensions.

Thus, for example in electrodynamics the letter $\mathbf{F}$ in the mathematical component is denoted by the word "force" and mapped onto the quantity which can be measured by a real spring balance or determined from the motions of masses. The mappings onto the basic domain establish only mappings onto the domain of reality of the forerunner theories. The forerunner theories are in this case mechanics and thermodynamics. For the description and prediction of experimental results in electrodynamics, these mappings are completely sufficient. The measurable quantities arise from the *domain of reality* of the forerunner theories. Although there are quantities in the mathematical component of the theory such as the symbol $\mathbf{j}$ which we refer to as the "electrical current density", for an experimental statement, however, it suffices that we can derive from the theory the fact that a wire is heated. We can use this fact to measure currents. It is not necessary to presume for this purpose that there are electrical currents "in reality" and that they somehow "flow" through wires. The word "current" initially serves only to be able to communicate more rapidly about parts of the theory.

Interpretation An *interpretation* of a physical theory is the specification of mapping principles by which some abstract computational quantities of the mathematical component MC

are given a *physical meaning* in relating them to parts of physical reality.[6] For some of the mathematical symbols, physical referents are defined. *In this sense, a physical theory is a partially-interpreted formal system.* It should be strictly distinguished if changes in or extensions of the mapping principles for the same MC are made, or on the other hand within an *alternative theory* the MC is also modified and, for example, different field equations are postulated.

The interpretation thus far described, which refers to the domain of reality of the forerunner theories, we shall call the *minimal interpretation*. It contains the minimal inventory of mapping principles which are required to attain a connection between the mathematical component of the theory and the observational level. On the basis of minimal interpretations, one can decide upon the empirical correctness of a physical theory. Further-reaching elements of an interpretation can neither be experimentally confirmed nor rejected. Many consistent extensions of the minimal interpretation and thus many interpretations can be envisioned. They are neither correct nor incorrect; it is the question of the explanatory power and comprehensibility of a theory which can provide a motivation to go beyond the minimal interpretation. Extended interpretations can provide valuable impulses for new research programmes. The search for a theory of the quantum measurement process is an example. In some interpretations, this search is superfluous (cf. Chap. 15). We should also point out that the concept 'minimal interpretation' is used in different senses in the literature.

2.4.2 Developed Reality*

The example of electrodynamics already demonstrates that most physicists go beyond the minimal interpretation. In the mathematical component of electrodynamics, along with terms which have a correspondence in the values indicated by measuring devices, additional terms such as the electric field $\mathbf{E}$ or the current density $\mathbf{j}$ occur, for which a correspondence in reality is likewise assumed. The mapping principles which belong to these terms will be called *hypothetical correspondence rules*, HCR. The domain onto which they map is called the *developed domain of reality*, DDR (see Fig. 2.5). Usually, there is a consensus concerning the addition of the hypothetical correspondence rules to the theory and the consequent extension of the domain of *physical reality* beyond the basic domain. We shall call this consensus with regard to quantum theory the *standard interpretation*. This term is also employed in differing fashions in the literature.

As is usual in physics, we will not speak of a transition to a new theory when only different hypothetical correspondence rules are chosen or they are dispensed with entirely, while the mathematical component MC and the basic domain BD remain unchanged. Different HCR lead to differing interpretations. The theory in whose MC the Maxwell equations occupy the central position is usually denoted –independently of whether or not reality is attributed to the electric field– as the 'Maxwell theory'. Which statements about reality are made by the physicist is left to the individual. A theory different from the Maxwell theory would only then be present if e. g. the Maxwell equations were changed or completely replaced.

[6] An interpretation is sometimes incorrectly denoted as the "philosophy" employed. Interpretations are however indispensable parts of the physical theories themselves.

* The sections marked with an asterisk * can be skipped over in a first reading.

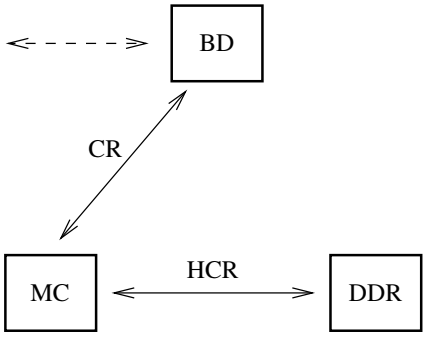


Figure 2.5: The correspondence principles relate the mathematical component MC of a physical theory to the basic domain BD of physical reality. Hypothetical correspondence rules HCR link the MC to a developed domain of reality DDR. The basic domain BD is the domain of reality of the forerunner theories. The dashed arrow indicates the direction in which the associated mathematical component is joined on.

It should have become clear that the determination of a developed domain of reality contains an element of arbitrariness and convention and that at least in part, reality becomes dependent on the theory. For yet another reason, the physical world comes into being with the aid of theory: *alternative theories*, which justify the same experiences (experimental results) in a different manner (i. e. with a different mathematical component MC), can in this sense correspond to different realities. Examples of this are on the one hand a theory of gravitation in terms of special relativity in a flat space-time, which uses the concept of a gravitational field, and on the other, the General Theory of Relativity in which the gravitational field is completely eliminated and space-time is curved. The standard interpretation in each case gives a different answer to the question of the existence of a gravitational field.

Electrodynamics furthermore shows that even within the standard interpretation, by no means is a correspondence in reality claimed for all mathematical quantities. For example, no reality is usually attributed to the gauge-dependent vector potential at a certain location in space. Vector potentials are considered to be only computational aids. In the Coulomb gauge, the vector potential changes instantaneously. There is however no real propagation process with a velocity greater than that of light connected with this.

2.5 Interpretations of Quantum Theory and Physical Reality*

2.5.1 The Minimal Interpretation*

How does quantum theory fit into the scheme described in the previous section? The *minimal interpretation* of quantum mechanics ascribes reality only to the preparations, transformations

*The sections marked with an asterisk * can be skipped over in a first reading.

and measurement devices. Only their reality exists and beyond them, there are no hypothetical correspondence rules HCR or developed domains of reality DDR. The mapping CR is performed within the classically describable reality domain (e. g. in terms of indicator readings of the measurement devices). The empirical knowledge within quantum mechanics can be formulated using the elements of classical physics. All other elements of the mathematical component MC of the quantum theory are only computational aids. This attitude can be accurately characterised by the intentionally exaggerated formulation of Niels Bohr: "*There is no quantum world.*"[7]. Electrons, atoms, etc. do not exist. This attitude is characterised in the theory of science as *instrumentalistic* and *pragmatic*. The advantage of limiting oneself to the minimal interpretation lies in the avoidance of apparent paradoxes. This is obtained at the price that no visualisation and hardly any physical intuition are stimulated.

In this interpretation, the single goal of quantum theory is to make precise predictions of possible results of measurements and to compute the probabilities of their occurrence. Further statements are superfluous and are not made[8]. Objectivity is guaranteed. After completion of the measurement, an observer can only read off the result but can no longer influence it. It is a part of the classical world. The empirical results thus obtained (e. g. via indications from measurement devices) can be described in the framework of classical physics as the associated forerunner theory. They can however not be explained or theoretically justified in terms of classical physics. For this purpose, one requires the mathematical component MC of quantum mechanics and some few correspondence rules. There are no unified statements in the literature concerning what precisely is contained in the *Copenhagen Interpretation* of the quantum theory. The minimal interpretation however certainly reflects the characteristic features of that interpretation.

2.5.2 The Standard Interpretation*

The *standard interpretation of the quantum theory*, which is generally accepted by physicists, goes somewhat further. It has been realised that individual quantum systems have properties which can be distinguished from the preparation of the states and their measurement; they thus do not describe relations between the system and the preparation or measurement apparatus. These physical properties include the electric charge, the baryon number, the mass and the magnitude of the spin of an elementary particle. They are referred to as *classical observables*[9]. Since they refer in part to the forerunner theories of quantum mechanics, it is reasonable to ascribe an objective reality to these properties and thus at the same time to the associated quantum system. *Quantum systems then have real media in a developed reality, which can*

[7] "There is no quantum world. There is only an abstract quantum-physical description. It is wrong to think that the task of physics is to find out how nature *is*. Physics concerns itself with what we can *say* about Nature." After [Pet 63, p. 12].

[8] Once again Niels Bohr: "In our description of nature, the purpose is not to disclose the real essence of the phenomena, but only to track down, so far as it is possible, relations between the manifold aspects of our experience." [Boh 85]

[9] The states of quantum systems which differ in these properties cannot be superposed. Linear combinations of such states are not physically relevant (*super selection rules*).

be called quantum objects. One can speak of individual quantum objects such as atoms, electrons, etc. and the following concept is associated with them in the standard interpretation:

The macroscopic effects of the measurement apparatus are described classically, e. g. in the form of the indicator deflections of the apparatus. These effects appear in this interpretation so to speak only on the "surface". They however can be attributed to the effects of quantum systems or quantum objects which are present in reality "beneath the surface". *There is a quantum world.* To be sure, we cannot perceive it with our senses; through them, we remain in the classical world. We act on the quantum world via the interfaces of the preparation and transformation apparatus, and register the reactions from that world once again in our classical world. The postulates in their formulation in Sect. 2.1.2 already reflect the correspondence rules of the standard interpretation .

Again, we do not claim that all the terms in the mathematical component MC correspond to elements of reality. *For the quantum state vector (or the density operator) which is determined by a preparation procedure, there is no correspondence to reality in the form of a physical object or physical properties. The correspondence can at most lie within the preparation process itself (cf. Sect. 2.1.2).* The state vector, as a mathematical symbol, allows the computation of the probability distributions in coaction with the measurement operators. In this sense, it is similar to the vector potential in electrodynamics. What evolves deterministically under the influence of a transformation apparatus are the predictable probability distributions of measurable results. The time-development operator $U(t, t_0)$ represents this evolution.

Even if this is not always explicitly mentioned in the following chapters, it will be useful for understanding the material to make clear to oneself whether a statement refers to the complete standard interpretation or is restricted to a more limited interpretation.

Further interpretations The *many-worlds interpretation* of the quantum theory is an example of an alternative interpretation which arrives at different statements on reality although retaining the mathematical component. We will deal briefly with it in Sect. 15.5. In the *subjectivistic interpretation* of the quantum theory, the state vector is interpreted as a statement about knowledge. Persons who have different knowledge then employ different state vectors. State vectors describe "beliefs" or "gambling commitments" [Fuc 03]. In this approach, evidently the measurement problem no longer appears. The modification of the state vector via the measurement is only an update of the knowledge of the observer through reading off the results of the measurement. With regard to the dynamic evolution between preparation and measurement, however, we must then ask the question as to why the knowledge of a person in the Schrödinger representation should change according to the time-dependent Schrödinger equation. Since knowledge and information are closely connected, this approach suggests that it would be reasonable to press on with the design of a programme to base the foundations of quantum theory solely on the concepts of information theory (see for example also [Fuc 03] or [Bub 05a]). The extreme position according to which reality comes into being by asking 'yes-no' questions and registering the answers "bit by bit" is taken in [Whe 90].

We will treat further interpretations and modifications of the quantum theory in Sects. 2.6, 10.8, 15.5, and 15.6.

2.6 Complementary Topics and Further Reading

- An elementary introduction to the physics of the entangled quantum world: [Aud 06] (book), [Aud 06a] (two articles).
- The two-slit experiment with atoms: [CM 91]. Review article on atom interferometry and atomic optics: [Ber 97]. Diffraction experiments with neutrons: [RW 00].
- Quantum objects exhibit mass-dependent interference effects in a gravitational field. They play a role in the axiomatic justification of the Riemannian structure of space-time [AL 91].
- The question of what happens to a quantum system between preparation and measurement is also the subject of considerations on *delayed choice*: [Whe 78], [Aul 00, Chap. 26].
- In this book, we consider only inertial systems. Whether or not quantum objects are present is a statement which depends on the state of acceleration of the observer (see Sect. 15.6)
- The Zeno effect: [NPN 97, p. 172], [Hom 97, Chap. 6].
- References to experiments on the Zeno effect: [IHB 90]; literature on the discussion of this experiment and suggestions for additional experiments are to be found in [NPN 97, p. 177].
- The philosophy of science and natural philosophy: [Sch 64], [Bal 70], [Mit 96], [Mai 96], [Hom 97], [Bub 97], [Mut 98], [Lal 01], [Esf 02].
- Forerunner theories, theoretical terms and correspondence rules: [Lud 83], [Lud 85], [Sch 90], [Sch 96a].
- Interpretations of the quantum theory: [Lud 55], [Pri 81], [Lud 89], [Lud 90], [Omn 94], [Lud 96], [FP 00]. Review article: [Lal 01].
- There are also other approaches to an axiomatic quantum theory which differ notably from that described here: the algebraic approach (for a review, see [Pri 81]) and the quantum-logical approach (a description can be found in [Mit 78]).
- We limit our considerations to quantum states in finite-dimensional Hilbert spaces. Among them, the states in two-dimensional spaces (qubits) play a particularly important role, since most investigations and applications of entanglement were initially carried out or implemented for these simplest states. At the same time, qubits represent the simplest generalisation of classical bits in view of quantum-information applications. For experimental implementations, e. g. with quantum-optical systems, states in infinite-dimensional Hilbert spaces however play an increasingly strong role. An overview of the theory and applications of *entanglement with continuous variables* can be found in [vLo 02] and [BvL 05].

- The goal of obtaining in an axiomatic manner the mathematical structure of the quantum theory (Hilbert-space formalism) from the macroscopic behaviour of preparation apparatus and measurement apparatus is pursued in [Lud 83] and [Lud 85].

© 2007 WILEY-VCH Verlag GmbH & Co.

3 The Simplest Quantum Systems: Qubits

All quantum systems which have no more than two linearly-independent states can be described by vectors in a two-dimensional Hilbert space $\mathcal{H}_2$. They are the simplest non-trivial quantum systems. Quantum states in $\mathcal{H}_2$ are called *qubits*, with a view to their later applications in quantum information theory. They have the form

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle \ , \quad |c_0|^2 + |c_1|^2 = 1 \ , \tag{3.1}$$

with the ONB $\{|0\rangle, |1\rangle\}$, which is also referred to as the *computational basis* or *standard basis*. Observables have at most two measurable values. The *qubit systems* are often also abbreviated simply as *qubits*.

Important physical implementations of qubit systems are:

- 2-level atoms (also atoms with more levels, when only two levels play a role in a particular process), and ions with two energy levels;
- Polarisation of spin-$\frac{1}{2}$ particles;
- Polarisation of single photons (horizontal $\leftrightarrow$ vertical or left-hand circular $\leftrightarrow$ right-hand circular);
- Ray paths in a two-path interferometer containing exactly one photon;
- Quantum dots;
- Modes of the electromagnetic field in a cavity resonator.

There are other qubit systems. The two-slit experiment and the Stern-Gerlach experiment can be similarly described in a simplified manner.

In Sect. 1.2.1, we met up with the operator basis of the transition operators. For calculations within $\mathcal{H}_2$, the operator basis consisting of the unit operator $\mathbb{1}$ and the σ *operators* (*Pauli operators*) is an important computational tool. Pauli operators are usually introduced in connection with the spin as an intrinsic angular momentum. Since, however, they have in general no physical significance with respect to an angular momentum when applied to other qubit systems, we can initially ignore this aspect. We introduce the Pauli operators in Sect. 3.1 and describe an often-used intuitive representation of quantum states in $\mathcal{H}_2$ and their dynamics by making use of the Bloch sphere in Sect. 3.2.

Qubit systems are the carriers of one unit of quantum information. The processing of quantum information, like that of classical information, is described by elementary operations, the so-called logic circuits or *gates*. *Quantum gates* are unitary transformations or measurements

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

on $\mathcal{H}_2$. We describe unitary gates which act only on a single qubit in Sect. 3.4. Implementations of qubit systems and quantum gates are introduced in Sects. 3.5 through 3.7.

3.1 Pauli Operators

The operator basis We introduce three Hermitian operators on $\mathcal{H}_2$, the σ_k with $k = 1, 2, 3$ or $k = x, y, z$:

$$\sigma_k^\dagger = \sigma_k \ , \tag{3.2}$$

for which we require

$$\sigma_k^2 = \mathbb{1} \ . \tag{3.3}$$

Since we want to assume that $\sigma_k \neq \mathbb{1}$ and that the eigenvalues are real, it follows from this and their Hermiticity that every operator σ_k has the eigenvalues $(+1)$ and (-1). At the same time, it can be seen from the spectral decomposition that the σ_k are also unitary:

$$\sigma_k^\dagger = \sigma_k^{-1} \tag{3.4}$$

and traceless:

$$\mathrm{tr}[\sigma_k] = 0 \ . \tag{3.5}$$

Figure 1.2 makes it clear how special the σ operators are. Figure 1.2 also shows that one could place the requirements of Hermiticity, unitarity and $\sigma \neq \mathbb{1}$ in the foreground. The eigenvalues $+1$ and -1 as well as Eqs. (3.3) and (3.5) would then be the consequences.

We wish to expand the σ_k into an operator basis through a requirement which relates them to one another (see Sect. 1.2.1). We can do this if in addition to Eqs. (3.2) and (3.3) we also ensure that the condition

$$\mathrm{tr}[\sigma_i \sigma_j] = 2\delta_{ij} \tag{3.6}$$

is fulfilled, by requiring of the anti-commutator ($[A, B]_+ := AB + BA$) that

$$[\sigma_i, \sigma_j]_+ = 2\delta_{ij}\mathbb{1} \ . \tag{3.7}$$

The operators $\{\frac{1}{\sqrt{2}}\mathbb{1}, \frac{1}{\sqrt{2}}\sigma_k\}$ *then form an orthonormal operator basis in Liouville space with respect to the scalar product (1.80). Every linear operator* A *can be represented in the form*

$$A = \frac{1}{2}\mathrm{tr}[A]\mathbb{1} + \frac{1}{2}\sum_{k=1}^{3}\mathrm{tr}[A\sigma_k]\sigma_k \ . \tag{3.8}$$

Angular-momentum operators The second important property of the σ_k is obtained through the following condition on the commutator ($[A,B]_- := AB - BA$), which complements Eq. (3.7):

$$[\sigma_i, \sigma_j]_- = 2i \sum_{k=1}^{3} \epsilon_{ijk} \sigma_k \ . \tag{3.9}$$

ϵ_{ijk} is a tensor which is totally antisymmetric in all indices, with $\epsilon_{123} = 1$. With this condition, the σ_k become formally proportional to angular-momentum operators and can describe observables of the 2-level system *spin*. In summary, we write for Eqs. (3.7) and (3.9)

$$\sigma_i \sigma_j = \delta_{ij} \mathbb{1} + i \sum_{k=1}^{3} \epsilon_{ijk} \sigma_k \ . \tag{3.10}$$

The Hermitian operators σ_k which obey Eq. (3.10) are called the *Pauli operators* or *σ-operators*. The equations show that the indices enter in a completely equivalent manner. This can be useful in some computations.

As a basis for an index-free formulation, we introduce the vector Pauli operator $\boldsymbol{\sigma} := \sigma_x \mathbf{e}_x + \sigma_y \mathbf{e}_y + \sigma_z \mathbf{e}_z$. The vectors $\mathbf{e}_x$, $\mathbf{e}_y$ and $\mathbf{e}_z$ are the Cartesian basis vectors of $\mathbb{R}^3$. Making use of Eq. (3.10), we can derive

$$(\boldsymbol{\sigma}\mathbf{a})(\boldsymbol{\sigma}\mathbf{b}) = (\mathbf{ab})\mathbb{1} + i\boldsymbol{\sigma}(\mathbf{a} \times \mathbf{b}) \tag{3.11}$$

for arbitrary vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$. This equation reduces to

$$(\boldsymbol{\sigma}\mathbf{e})(\boldsymbol{\sigma}\mathbf{e}) = \mathbb{1} \tag{3.12}$$

for arbitrary normalised vectors $\mathbf{e}$. We thus obtain with the aid of the spectral decomposition for the expansion of the exponential function (cf. Eq. (1.50)):

$$\exp(i\theta \mathbf{e}\boldsymbol{\sigma}) = \mathbb{1} + i\theta \mathbf{e}\boldsymbol{\sigma} - \frac{1}{2}\theta^2 \underbrace{(\mathbf{e}\boldsymbol{\sigma})^2}_{=\mathbb{1}} + \frac{i}{3!}\theta^3 \boldsymbol{e\sigma} \pm \dots \ , \tag{3.13}$$

and, after collecting terms, the frequently-used relation

$$\exp(i\theta \mathbf{e}\boldsymbol{\sigma}) = (\cos\theta)\mathbb{1} + i(\sin\theta)\mathbf{e}\boldsymbol{\sigma} \ . \tag{3.14}$$

The matrix representation If we take the ONB of the eigenvectors $|0\rangle$ and $|1\rangle$ of σ_z, with the eigenvalues $(+1)$ or (-1), respectively,

$$\begin{aligned} \sigma_z|0\rangle &= +|0\rangle \\ \sigma_z|1\rangle &= -|1\rangle \end{aligned} \tag{3.15}$$

as the computational basis, then the matrix representation of σ_z is

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \ . \tag{3.16}$$

Making use of Eqs. (3.2) and (3.10), the matrix representations of σ_x and σ_y can be calculated individually in the computational basis. Here, we just give the result. We denote these *Pauli matrices* likewise by the same symbol as the operators and give directly their dyadic decompositions:

$$\begin{aligned}
\sigma_x &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = |0\rangle\langle 1| + |1\rangle\langle 0| \\
\sigma_y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = -i\big(|0\rangle\langle 1| - |1\rangle\langle 0|\big) \\
\sigma_z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = |0\rangle\langle 0| - |1\rangle\langle 1| \,.
\end{aligned} \tag{3.17}$$

It is useful to note the action of the σ operators on the vectors of the computational basis:

$$\begin{aligned}
&\sigma_x|0\rangle = |1\rangle && \sigma_x \text{ exchanges (\textit{bit flip})} \\
&\sigma_x|1\rangle = |0\rangle \\
&\sigma_y|0\rangle = i|1\rangle && \sigma_y \text{ exchanges and introduces the phase } \pm i \\
&\sigma_y|1\rangle = -i|0\rangle \\
&\sigma_z|0\rangle = +|0\rangle && \sigma_z \text{ introduces the phase } \pm 1 \text{ (\textit{phase flip})}. \\
&\sigma_z|1\rangle = -|1\rangle.
\end{aligned} \tag{3.18}$$

With these relations, we can finally verify the orthonormal eigenstates of σ_x and σ_y directly as

$$\begin{aligned}
\sigma_x|0_x\rangle &= |0_x\rangle & \sigma_x|1_x\rangle &= -|1_x\rangle \\
|0_x\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) & |1_x\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)
\end{aligned} \tag{3.19}$$

and

$$\begin{aligned}
\sigma_y|0_y\rangle &= |0_y\rangle & \sigma_y|1_y\rangle &= -|1_y\rangle \\
|0_y\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) & |1_y\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle)
\end{aligned} \tag{3.20}$$

in the computational basis.

3.2 Visualisation of Qubits on the Bloch Sphere

We first carry out some preparatory mathematical considerations based on the decomposition in the operator basis, to which we shall refer frequently in later sections. An operator ρ with

the properties $\rho^\dagger = \rho$ and $\mathrm{tr}[\rho] = 1$ can be written in the following form with the aid of Eq. (3.8):

$$\rho = \frac{1}{2}(\mathbb{1} + \mathbf{r}\boldsymbol{\sigma}) \tag{3.21}$$

with a real vector $\mathbf{r}$

$$\mathbf{r} := \mathrm{tr}[\rho\boldsymbol{\sigma}] \; . \tag{3.22}$$

Using Eqs. (3.5) and (3.10), we obtain

$$\begin{aligned} \mathrm{tr}[\rho^2] &= \frac{1}{4}\mathrm{tr}[\mathbb{1} + 2\mathbf{r}\boldsymbol{\sigma} + \sum_{i,j} r_i r_j \sigma_i \sigma_j] \\ &= \frac{1}{2}(1 + |\mathbf{r}|^2) \; . \end{aligned} \tag{3.23}$$

As a special case, we consider the operator $\rho := |\psi\rangle\langle\psi|$ which is constructed with an arbitrary normalised vector $|\psi\rangle \in \mathcal{H}_2$. For this operator, the conditions $\rho^2 = \rho$ and $\mathrm{tr}[\rho^2] = \mathrm{tr}[\rho] = 1$ are valid, and it follows with (3.23) that $|\mathbf{r}| = 1$. Inserting ρ into (3.22), we obtain as an interpretation of $\mathbf{r}$ the expectation value

$$\mathbf{r} = \langle\psi|\boldsymbol{\sigma}|\psi\rangle \; . \tag{3.24}$$

With $\langle\psi|\mathbf{r}\boldsymbol{\sigma}|\psi\rangle = \mathbf{r}\mathbf{r} = 1$, we find the additional result

$$\mathbf{r}\boldsymbol{\sigma}|\psi\rangle = |\psi\rangle \; . \tag{3.25}$$

Here, we have used the fact that $\boldsymbol{r\sigma}|\psi\rangle$ is a normalised vector as a consequence of Eq. (3.12). We have thus found two results: *For an arbitrary normalised state vector $|\psi\rangle \in \mathcal{H}_2$, the expectation value of the vector Pauli operator $\boldsymbol{\sigma}$ is a real vector $\mathbf{r} \in \mathbb{R}^3$ with the magnitude* 1. *At the same time, $|\psi\rangle$ is an eigenvector of the operator $\mathbf{r}\boldsymbol{\sigma}$, with the eigenvalue* $(+1)$.

The Bloch sphere Via the expectation value of $\boldsymbol{\sigma}$, each qubit $|\psi\rangle$ is, according to Eq. (3.24), uniquely associated with a vector $\mathbf{r}$ in $\mathbb{R}^3$; it is called the *Bloch vector*. Its endpoint lies on the surface of a unit sphere, the *Bloch sphere* (compare Fig. 3.1). We call this endpoint the *Bloch point*. With the aid of $\mathbf{r}$, we can visualise the state vectors $|\psi\rangle$ as well as the effects of the measurement dynamics and the unitary dynamics on $|\psi\rangle$ three-dimensionally in a simple way. Therein lies the importance of the Bloch vectors.

$\mathbf{r}\boldsymbol{\sigma}$ is Hermitian. $\mathbf{r}$ is normalised and Eq. (3.12) correspondingly applies. We can therefore adopt the arguments from the beginning of this chapter: the eigenvalues of the operators $\mathbf{r}\boldsymbol{\sigma}$ are $(+1)$ and (-1). The eigenvector with the eigenvalue $(+1)$ is $|\psi\rangle$, with the Bloch vector $\mathbf{r}$ from Eq. (3.24). The second eigenvector of $\mathbf{r}\boldsymbol{\sigma}$ is perpendicular to $|\psi\rangle$ ($\langle\chi|\psi\rangle = 0$) and obeys

$$\mathbf{r}\boldsymbol{\sigma}|\chi\rangle = -|\chi\rangle \; . \tag{3.26}$$

Multiplication by $\langle\chi|$ leads using $|\langle\chi|\boldsymbol{\sigma}|\chi\rangle| = +1$ (magnitude of the Bloch vector of $|\chi\rangle$) to

$$\langle\chi|\boldsymbol{\sigma}|\chi\rangle = -\mathbf{r} \; . \tag{3.27}$$

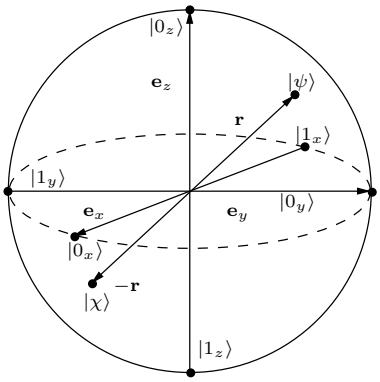


Figure 3.1: The Bloch sphere with the Bloch vectors associated to the eigenstates of the Pauli operators. The Bloch vectors $\mathbf{r}$ and $-\mathbf{r}$, corresponding to two arbitrary orthonormal qubits $|\psi\rangle$ and $|\chi\rangle$, are mirror-symmetric with respect to the coordinate origin.

The Bloch vector with the eigenvalue (-1) is formed by reflection of $\mathbf{r}$ through the centre of the sphere (cf. Fig. 3.1).

In general, two orthogonal qubits correspond to two Bloch vectors which are related to one another by a reflection through the coordinate origin. The associated operator is found directly via the Bloch vector $\mathbf{r}$ to be $\mathbf{r}\boldsymbol{\sigma}$. In particular, we find that the Bloch vectors corresponding to the eigenstates of the three Pauli operators for the eigenvalues $(+1)$ agree with the basis vectors $\mathbf{e}_x, \mathbf{e}_y$, and $\mathbf{e}_z$. For the eigenvalues (-1), the Bloch vectors point in the respective opposite directions. This is shown in Fig. 3.1. We still want to parametrise the Bloch points by making use of polar coordinates. An arbitrary qubit $|\psi\rangle$ (see Eq. (3.1)) can, by using two parameters θ and φ, always be represented in the form

$$\begin{aligned} |\psi\rangle &= e^{-i\frac{\varphi}{2}} \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\frac{\varphi}{2}} \sin\left(\frac{\theta}{2}\right)|1\rangle \\ &= e^{-i\frac{\varphi}{2}} \left\{\cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi} \sin\left(\frac{\theta}{2}\right)|1\rangle\right\} \end{aligned} \tag{3.28}$$

($0 \leq \varphi \leq 2\pi$, $0 \leq \theta \leq \pi$). With Eqs. (3.17) and (3.24), we can then readily compute the Bloch vector associated with $|\psi\rangle$

$$\mathbf{r} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)\ . \tag{3.29}$$

In the visualisation on the Bloch sphere (Fig. 3.2), the parameters θ and φ of Eq. (3.28) thus take on the role of polar coordinates of the Bloch point . It is typical that in connection with the associated state vector $|\psi\rangle$, half polar angles appear.

This has as a result that on the one hand, for a given qubit $|\psi\rangle$, the Bloch vector $\mathbf{r}$ is well determined, but that on the other, a Bloch vector is associated with two different qubits. If we for example rotate the Bloch vector around the z-axis into the y–z plane ($\varphi = \frac{\pi}{2}$), then it returns after a 2π rotation ($\theta = 0 \to \theta = \pi \to \theta = 0$) to its initial position. Equation (3.28)

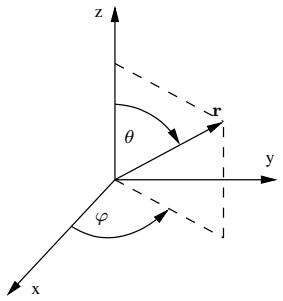


Figure 3.2: Polar coordinates allow us to visualise the angles θ and φ of Eq. 3.28.

shows that the associated state vector changes its sign in this process, $|\psi\rangle \rightarrow -|\psi\rangle$. Only after a full 4π rotation of the Bloch vector does the state vector again return to its initial position, $|\psi\rangle \rightarrow +|\psi\rangle$.

It is useful to note the components of the operator $\mathbf{r}\boldsymbol{\sigma}$ in the computational basis:

$$\mathbf{r}\boldsymbol{\sigma} = \begin{pmatrix} \cos\theta & e^{-i\varphi}\sin\theta \\ e^{i\varphi}\sin\theta & -\cos\theta \end{pmatrix} . \tag{3.30}$$

We denote the eigenvectors of $\mathbf{r}\boldsymbol{\sigma}$ by $|0_{\boldsymbol{r}}\rangle$ and $|1_{\boldsymbol{r}}\rangle$. The eigenvector $|0_{\boldsymbol{r}}\rangle = |\psi\rangle$ was already defined in Eq. (3.28). We note also the eigenvector $|1_{\mathbf{r}}\rangle = |\chi\rangle$ belonging to the eigenvalue -1:

$$|0_{\boldsymbol{r}}\rangle = |\psi\rangle = \begin{pmatrix} e^{-i\varphi/2}\cos(\theta/2) \\ e^{i\varphi/2}\sin(\theta/2) \end{pmatrix} ; \quad |1_{\boldsymbol{r}}\rangle = |\chi\rangle = \begin{pmatrix} -e^{-i\varphi/2}\sin(\theta/2) \\ e^{i\varphi/2}\cos(\theta/2) \end{pmatrix} . \tag{3.31}$$

3.3 Visualisation of the Measurement Dynamics and the Unitary Dynamics

Projective measurements In the postulates, we introduced two types of dynamics, the measurement dynamics and the unitary dynamics. The measurement dynamics of projective measurements can be quite simply visualised. An observable operator in the two-dimensional Hilbert space $\mathcal{H}_2$ of the qubits has two orthogonal eigenvectors, $|0\rangle$ and $|1\rangle$. We can always interpret them as eigenvectors of σ_z. On the Bloch sphere, we fix for this the direction of $\mathbf{e}_z$ in the direction of the Bloch vector of $|0\rangle$. A measurement of this observable then causes a quantum transition from an initial state $|\psi\rangle$ into the final state $|0\rangle$ or $|1\rangle$, depending on the result of the measurement (see Fig. 3.3).

If we write $|\psi\rangle$ in the form

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle , \tag{3.32}$$

then the corresponding probabilities are $|c_0|^2$ and $|c_1|^2$. The projection r_z of the Bloch vector $\mathbf{r}$ onto the z axis, using Eq. (3.29), is

$$r_z = \cos\theta = \cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2} = |c_0|^2 - |c_1|^2 . \tag{3.33}$$

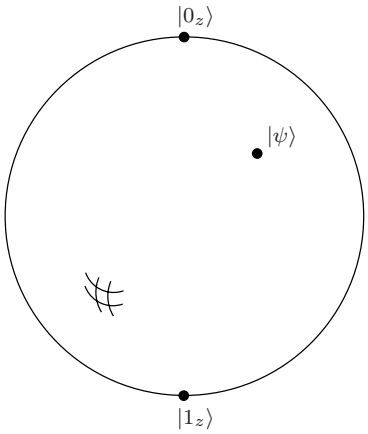


Figure 3.3: In a projective measurement of an observable operator with the eigenvectors $|0_z\rangle$ and $|1_z\rangle$, the state vector $|\psi\rangle$ makes a jump into $|0_z\rangle$ or $|1_z\rangle$, depending on the result of the measurement. The corresponding Bloch points on the surface of the Bloch sphere are shown.

Since $|c_0|^2 + |c_1|^2 = 1$, it follows for the probability $|c_0|^2$ that

$$|c_0|^2 = \frac{r_z + 1}{2} \,. \tag{3.34}$$

Unitary transformations We want to complete this section by showing a visualisation of the unitary dynamics, which transform an initial state $|\psi\rangle$ with a unitary operator U according to

$$|\psi'\rangle = U|\psi\rangle \tag{3.35}$$

into the final state $|\psi'\rangle$. To this end, we start from some mathematical properties of unitary transformations and their matrix representations in $\mathcal{H}_2$.

Both the rows and the columns of a unitary matrix are pairwise orthonormal with respect to one another ($\sum_j U_{ij}^* U_{kj} = \delta_{ik}$). The evaluation of the corresponding relations, which we shall not elaborate here, shows that for a unitary matrix U in $\mathcal{H}_2$, four real parameters κ, λ, μ and ν exist, so that U can be written as a matrix of the form

$$U = e^{i\kappa} \begin{pmatrix} e^{-i\lambda/2}\cos(\mu/2)e^{-i\nu/2} & -e^{-i\lambda/2}\sin(\mu/2)e^{i\nu/2} \\ e^{i\lambda/2}\sin(\mu/2)e^{-i\nu/2} & e^{i\lambda/2}\cos(\mu/2)e^{i\nu/2} \end{pmatrix} \tag{3.36}$$

or as an operator product

$$U = \exp(i\kappa)\exp\left(-\frac{i}{2}\lambda\sigma_z\right)\exp\left(-\frac{i}{2}\mu\sigma_y\right)\exp\left(-\frac{i}{2}\nu\sigma_z\right) \,. \tag{3.37}$$

Only the Pauli operators σ_y and σ_z occur in these expressions. It is helpful to introduce the unitary operator $\hat{U}$:

$$\hat{U} := e^{-i\kappa}U \,, \tag{3.38}$$

which is the same as U except for the global phase factor $e^{i\kappa}$. From Eq. (3.36), we can directly read off the following properties of the matrix representation of $\hat{U}$:

$$\begin{aligned} \hat{U}_{00} &= \hat{U}_{11}^* \\ \hat{U}_{10} &= -\hat{U}_{01}^* \\ \hat{U}_{00}\hat{U}_{11} - \hat{U}_{01}\hat{U}_{10} &= 1 \,. \end{aligned} \tag{3.39}$$

We decompose $\hat{U}$ in terms of the operator basis $\{\frac{1}{\sqrt{2}}\mathbb{1}, \frac{1}{\sqrt{2}}\sigma_k\}$ according to Eq. (3.8)

$$\hat{U} = v_0\mathbb{1} - i\boldsymbol{v}\boldsymbol{\sigma} \,. \tag{3.40}$$

where v_0 and $\boldsymbol{v}$ are determined by taking the trace:

$$v_0 = \frac{1}{2}\mathrm{tr}[\hat{U}] \,, \quad \boldsymbol{v} = \frac{i}{2}\mathrm{tr}[\hat{U}\boldsymbol{\sigma}] \,. \tag{3.41}$$

Making use of the relations (3.39), one can readily show that v_0 and $\boldsymbol{v}$ are real. The unitarity relation $\hat{U}^\dagger\hat{U} = \mathbb{1}$ is, with Eq. (3.11), equivalent to the condition

$$v_0^2 + \boldsymbol{v}\boldsymbol{v} = 1 \,. \tag{3.42}$$

It confirms again that a unitary 2×2 matrix is determined by a global phase and three real parameters.

Since the condition (3.42) is fulfilled, Eq. (3.40) has the structure of Eq. (3.14). For later use, we introduce the angle ϕ and the unit vector $\boldsymbol{e}$ according to

$$v_0 := \cos\frac{\phi}{2} \,, \quad \boldsymbol{v} = \left(\sin\frac{\phi}{2}\right)\mathbf{e} \,. \tag{3.43}$$

For a given $\hat{U}$, the quantities ϕ and $\boldsymbol{e}$ are determined by Eq. (3.41). With Eq. (3.14), Eq. (3.40) takes on the form

$$\hat{U} = \exp\left(-i\frac{\phi}{2}\boldsymbol{e}\boldsymbol{\sigma}\right) \,. \tag{3.44}$$

Every unitary transformation in $\mathcal{H}_2$ can be written uniquely in the form (3.44) up to a phase $e^{i\kappa}$.

Finally, we want to explain the effects of U on the Bloch sphere. The phase factor $e^{i\kappa}$ has no effect on the Bloch vector. Due to the symmetry of the representation, we can fix $\mathbf{e}_z$ in the direction of $\mathbf{e}$ without loss of generality. The operator $\hat{U}$ then – with Eq. (3.44) – takes on the form

$$\hat{U} = e^{-i\frac{\phi}{2}}|0\rangle\langle 0| + e^{+i\frac{\phi}{2}}|1\rangle\langle 1| \tag{3.45}$$

(compare Eq. (1.52)). Making use of Eq. (3.28), we can directly read off that $\hat{U}|\psi\rangle$ is obtained from $|\psi\rangle$ by the substitution $\varphi \to \varphi + \phi$. A simple interpretation of the effect of U on the

Bloch sphere follows: *When the state $|\psi\rangle$ is represented by the Bloch vector $\boldsymbol{r}$, then the Bloch vector $\boldsymbol{r}'$ associated with the state vector $|\psi'\rangle$ which has undergone a unitary transformation*

$$|\psi'\rangle = R_{\mathbf{e}}(\phi)|\psi\rangle, \quad R_{\mathbf{e}}(\phi) = \exp\left(-i\frac{\phi}{2}\boldsymbol{e}\boldsymbol{\sigma}\right) \tag{3.46}$$

is generated by a rotation of $\boldsymbol{r}$ by an angle ϕ on a cone around the axis $\boldsymbol{e}$ (cf. Fig. (3.4)). The quantities $\boldsymbol{e}$ and ϕ are determined by Eqs. (3.41) and (3.43). The unitary transformation $\exp(-i\frac{\phi}{2}\sigma_z)$ (or $\exp(-i\frac{\phi}{2}\sigma_x)$ or $\exp(-i\frac{\phi}{2}\sigma_y)$) corresponds to a rotation of the Bloch vector around the z axis (or the x axis or the y axis, respectively) by the angle ϕ.

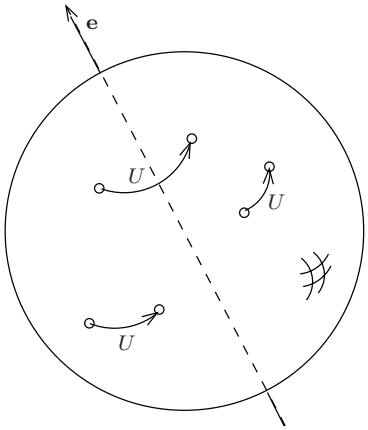


Figure 3.4: The effect of the unitary transformation U on the end points of the Bloch vectors of pure states. The vector $\mathbf{e}$ is given by Eqs. (3.41) and (3.43).

Example: Rabi oscillations We wish to describe a physical situation in which the Bloch vector undergoes periodic motion. A single photon within a cavity oscillator interacts with a single 2-level atom whose energy levels are denoted by $|0\rangle$ and $|1\rangle$. For simplicity, we consider the case of resonance; i.e. the energy difference between the two levels is equal to the photon energy. Then the photon will be periodically absorbed and emitted. The probability $|c_0|^2$ of finding the atom in its excited state $|0\rangle$ oscillates in time. The corresponding frequency is called the *Rabi frequency* Ω_R. Its value is a measure of the strength of the interactions between the atom and the quantised radiation field. A quantum-electrodynamics calculation shows that the resulting influence on the state vector of the atom can be described by the unitary transformation

$$U(t, t_0) = \exp\left[-i(t - t_0)\frac{\Omega_R}{2}\sigma_x\right] . \tag{3.47}$$

As we have seen, this has the result that the Bloch vector $\mathbf{r}$ rotates with the frequency Ω_R on a cone around the x axis (compare Fig. 3.5). In the limiting case of a completely

open cone, it rotates in the y–z plane. The probability $|c_0|^2$ is found from the projection $r_z(t) = r_z(t - \frac{2\pi}{\Omega_R})$ of $\mathbf{r}$ onto the z axis, giving

$$|c_0|^2 = \frac{r_z(t) + 1}{2} \,. \tag{3.48}$$

The maximum value of the projection $r_z(t)$ can attain its largest possible magnitude of 1 only when $\mathbf{r}$ rotates within the y–z plane. For this, it is sufficient that $\mathbf{r}$ be directed e. g. parallel to $\mathbf{e}_z$ or $-\mathbf{e}_z$ by an initial projection measurement. These correspond to the states $|0\rangle$ or $|1\rangle$. In the other limiting case, the atom is prepared in one of the eigenstates $|0_x\rangle$ or $|1_x\rangle$ of σ_x (cf. Fig. 3.1). Then the Bloch vector lies along $\mathbf{e}_x$ or $-\mathbf{e}_x$ and remains unchanged under the influence of the interactions. The reason for this is to be found in the quantum-mechanical computation: The Hamiltonian of the composite system, which is composed of the sum of the Hamiltonians for the free photon, the free atom, and for the interactions, has two eigenstates. When the composite system is in one of these eigenstates, it remains in that state. The corresponding atomic states are found to be $|0_x\rangle$ and $|1_x\rangle$.

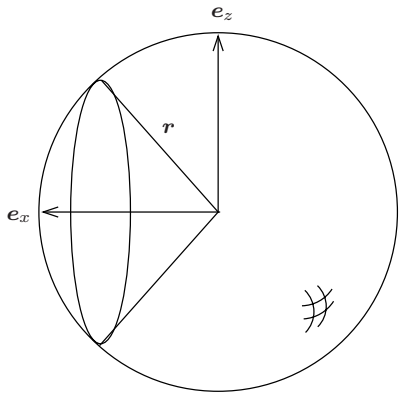


Figure 3.5: Rabi oscillations of the Bloch vector $\mathbf{r}$.

3.4 Quantum Gates for Single Qubit Systems

We wish to summarise the mathematically and physically especially relevant unitary transformations in $\mathcal{H}_2$. In the previous section, we saw that the transformations $\exp(i\phi\sigma_k)$ cause rotations of the Bloch vector around the coordinate axes $\boldsymbol{e}_k$. The essential point is that according to Eq. (3.37), every arbitrary unitary transformation can be obtained by repeated applications of these particular transformations.

NOT and $\sqrt{\text{NOT}}$ gates In quantum information theory, certain unitary operators play particular roles, and these are also denoted as *quantum gates*. Among them are the three Pauli

operators σ_k, whose effects we have already described. The Pauli operator σ_x is also termed a *NOT gate* (cf. Tab. 3.1). It exchanges $|0\rangle$ and $|1\rangle$

$$NOT := |0\rangle\langle 1| + |1\rangle\langle 0| = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad (NOT)(NOT)^\dagger = \mathbb{1} . \tag{3.49}$$

For the $\sqrt{NOT}$ *gate* (or square-root-of-not gate), we find

$$\sqrt{\text{NOT}} := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix} , \quad \sqrt{NOT}\sqrt{NOT} = iNOT . \tag{3.50}$$

The *phase gate*

$$\phi(\alpha) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{pmatrix} = |0\rangle\langle 0| + e^{i\alpha}|1\rangle\langle 1| , \tag{3.51}$$

causes a phase shift of the $|1\rangle$ component of a vector (*phase shifter*). $\phi(\alpha)$ can also be written in the form

$$\phi(\alpha) = e^{i\frac{\alpha}{2}} e^{-i\frac{\alpha}{2}\sigma_z} . \tag{3.52}$$

It therefore is identical to one of the rotations mentioned above, up to a global phase factor.

The Hadamard gate We furthermore introduce the unitary and Hermitian *Hadamard gate* H:

$$H := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}}(\sigma_x + \sigma_z) , \quad H^\dagger = H = H^{-1} , \quad H^2 = \mathbb{1} , \tag{3.53}$$

which can be written as a sum of Pauli operators. H is identical to its inverse and converts the vectors of the computational basis into the eigenvectors of σ_x (compare Eq. (3.19)):

$$\begin{aligned} H|0\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ H|1\rangle &= \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) . \end{aligned} \tag{3.54}$$

We also note the relations

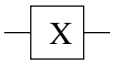
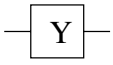
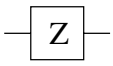
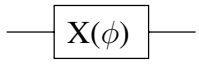
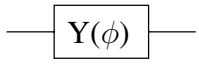
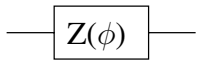
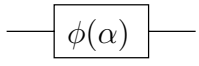
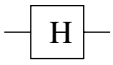
$$H\sigma_x H = \sigma_z , \qquad H\sigma_y H = -\sigma_y , \qquad H\sigma_z H = \sigma_x , \tag{3.55}$$

which can be obtained e. g. by applying the relation (3.7).

The question arises as to which is the axis of rotation $\mathbf{e}$ and through which angle of rotation θ the Bloch vector is turned when the Hadamard gate acts upon a state. We employ the results of Sect. 3.3. With the phase factor $e^{-i\kappa} = -i$, we obtain from H a unitary operator $\hat{U}$ (cf. Eq. (3.39)) which has the action

$$\hat{U}|0\rangle = \frac{-i}{\sqrt{2}}(|0\rangle + |1\rangle) , \quad \hat{U}|1\rangle = \frac{-i}{\sqrt{2}}(|0\rangle - |1\rangle) . \tag{3.56}$$

Table 3.1: Frequently-used gates for one qubit.

	Gate	Operator symbol	Operator	Matrix
σ_x operator		σ_x	$\|0\rangle\langle 1\| + \|1\rangle\langle 0\|$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
σ_y operator		σ_y	$-i(\|0\rangle\langle 1\| - \|1\rangle\langle 0\|)$	$\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$
σ_z operator		σ_z	$\|0\rangle\langle 0\| - \|1\rangle\langle 1\|$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
NOT		NOT, σ_x	$\|0\rangle\langle 1\| + \|1\rangle\langle 0\|$	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
$\sqrt{\text{NOT}}$		$\sqrt{\text{NOT}}$	$e^{i\frac{\pi}{4}\sigma_x}$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$
x rotation		$R_x(\phi)$	$e^{-i\frac{\phi}{2}\sigma_x}$	$\begin{pmatrix} \cos\frac{\phi}{2} & -i\sin\frac{\phi}{2} \\ -i\sin\frac{\phi}{2} & \cos\frac{\phi}{2} \end{pmatrix}$
y rotation		$R_y(\phi)$	$e^{-i\frac{\phi}{2}\sigma_y}$	$\begin{pmatrix} \cos\frac{\phi}{2} & -\sin\frac{\phi}{2} \\ \sin\frac{\phi}{2} & \cos\frac{\phi}{2} \end{pmatrix}$
z rotation		$R_z(\phi)$	$e^{-i\frac{\phi}{2}\sigma_z}$	$\begin{pmatrix} e^{-i\frac{\phi}{2}} & 0 \\ 0 & e^{i\frac{\phi}{2}} \end{pmatrix}$
Phase		$\phi(\alpha)$	$\|0\rangle\langle 0\| + e^{i\alpha}\|1\rangle\langle 1\|$	$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\alpha} \end{pmatrix}$
Hadamard		H	$\frac{1}{\sqrt{2}}(\sigma_x + \sigma_z)$ $= e^{-i\frac{\pi}{4}\sigma_y}\sigma_z$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

We can make use of Eq. (3.17) to readily evaluate Eq. (3.41) and obtain the following result: the Hadamard gate causes a rotation through the angle

$$\theta = 180^\circ \tag{3.57}$$

around the axis

$$\mathbf{e} = \frac{1}{\sqrt{2}}(e_x + e_z)\ . \tag{3.58}$$

Additional gates are listed in Tab. (3.1).

3.5 Spin-$\frac{1}{2}$

A particularly important implementation of a *qubit system* is the *spin* of quantum number $\frac{1}{2}$. It is an internal degree of freedom of elementary particles, such as electrons, and is described by means of a two-dimensional Hilbert space $\mathcal{H}_2$ (spin space). The associated observable is

$$\boldsymbol{S} = \frac{\hbar}{2}\boldsymbol{\sigma} \; . \tag{3.59}$$

Its components, according to Eq. (3.9), obey the commutation relations for angular momenta,

$$[S_i, S_j]_- = i\hbar\epsilon_{ijk}S_k \; . \tag{3.60}$$

A magnetic moment is associated with the spin, with the observable

$$\mathbf{M} = \gamma\mathbf{S} \; . \tag{3.61}$$

The gyromagnetic ratio γ for electrons has the value $\frac{e}{mc}$. In a magnetic field $\mathbf{B}$, the interaction between the magnetic field and the magnetic moment leads to a Hamiltonian

$$H = -\gamma\mathbf{B}\mathbf{S} \; . \tag{3.62}$$

We orient $\mathbf{e}_z$ in the direction of $\mathbf{B}$ and introduce $\omega := -\gamma B$ with $B = |\mathbf{B}|$. Then the Hamiltonian takes on the form $H = \frac{\hbar\omega}{2}\sigma_z$ with the eigenvalues $\pm\frac{\hbar\omega}{2}$ and the eigenstates $|0_z\rangle$ and $|1_z\rangle$. For a spin-$\frac{1}{2}$ system, the observable σ_z can, depending on the physical situation, be interpreted up to a factor as a component of the magnetic moment or as the energy of the system in the magnetic field (2-level system).

3.6 Photon Polarisations

In the case of linear polarisation, the monochromatic electromagnetic wave fields have the form

$$\boldsymbol{E}_H \sim \mathbf{e}_H \exp i(\boldsymbol{kr} - \omega t); \quad \boldsymbol{E}_V \sim \mathbf{e}_V \exp i(\boldsymbol{kr} - \omega t); \tag{3.63}$$

with the propagation vector $\boldsymbol{k}$ (see Fig. 3.6). The indices H and V refer to horizontal and vertical polarisations or oscillation planes. Another basis is given by

$$\mathbf{e}_{H'} = \mathbf{e}_{+45^\circ} = \frac{1}{\sqrt{2}}(\mathbf{e}_H + \mathbf{e}_V); \quad \mathbf{e}_{V'} = \mathbf{e}_{-45^\circ} = \frac{1}{\sqrt{2}}(\mathbf{e}_H - \mathbf{e}_V) \; . \tag{3.64}$$

It corresponds to a rotation of the plane of oscillation by an angle of 45° around the $\mathbf{k}$ axis. The right-hand and left-hand circularly polarised waves are given by

$$\boldsymbol{E}(R, L) \sim \mathbf{e}(R, L) \exp i(\boldsymbol{kr} - \omega t) \tag{3.65}$$

with

$$\mathbf{e}(R) = \frac{1}{\sqrt{2}}(\mathbf{e}_H + i\mathbf{e}_V); \quad \mathbf{e}(L) = \frac{1}{\sqrt{2}}(\mathbf{e}_H - i\mathbf{e}_V) \; . \tag{3.66}$$

The state vectors of the quantum system *photon* corresponding to these polarisations are vectors in a Hilbert space $\mathcal{H}_2$ and are thus qubits. The correspondence is given by:

$$\mathbf{e}_H \leftrightarrow |H\rangle = |0\rangle, \quad \mathbf{e}_V \leftrightarrow |V\rangle = |1\rangle \tag{3.67}$$

$$\mathbf{e}_{H'} \leftrightarrow |H'\rangle = |+45^\circ\rangle = \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle) = |0_x\rangle$$

$$\mathbf{e}_{V'} \leftrightarrow |V'\rangle = |-45^\circ\rangle = \frac{1}{\sqrt{2}}(|H\rangle - |V\rangle) = |1_x\rangle \tag{3.68}$$

$$\mathbf{e}_R \leftrightarrow |R\rangle = \frac{1}{\sqrt{2}}(|H\rangle + i|V\rangle) = |0_y\rangle$$

$$\mathbf{e}_L \leftrightarrow |L\rangle = \frac{1}{\sqrt{2}}(|H\rangle - i|V\rangle) = |1_y\rangle \tag{3.69}$$

The connection with the eigenvectors of the Pauli operators has been indicated for the three types of polarisation.

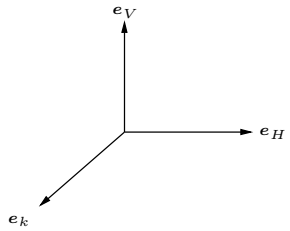


Figure 3.6: Polarisation vectors and the propagation vector (wavevector k) for linearly-polarised photons.

3.7 Single Photons in a Beam Splitter and in an Interferometer

We wish to introduce another qubit system which is especially significant for quantum-optical experiments relating to quantum information theory. It consists of a single photon on which a transformation apparatus in the sense of Sect. 1.2 acts; it is constructed from a series of individual beam splitters, phase shifters, and mirrors. Simple examples of such a photonic network for quantum information processing are beam splitters and interferometers themselves (cf. Figs. 3.7 and 3.8).

The photon is the quantum of an electromagnetic radiation field with well-determined mode functions. In our case, these are plane waves which are characterised by a wavevector. The optical setup which we consider is assumed to be so simple that only two photonic modes or paths, the 0 path and the 1 path, can be traced through it[1]. If the photon is registered before,

[1] If, for simplicity's sake, we speak of paths, then this should not be understood to imply that the photon "flies along this path" (compare Sect. 2.1.4). The transformation apparatus can also be used for classical light; the paths then correspond to classical light rays. This defines the paths. The illustrations are to be understood in this sense.

within, or after the apparatus, it is found on only one of the paths. Therefore, we can describe the photon in this situation as a qubit, whereby the states $|0\rangle$ or $|1\rangle$ refer to the two possible paths. The normalised photon state is of the form

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle \tag{3.70}$$

with $|c_0|^2 + |c_1|^2 = 1$. If we place a detector in the 0 path (0 detector), then it registers a signal with the probability $|c_0|^2$. A corresponding expression holds for the 1 path. The optical apparatus used have two entry paths and two exit paths, and their effects are described by a series of unitary transformations of the state vector $|\psi\rangle$. They modify the probability of registering the photon on a particular path, leaving the overall probability conserved.

Phase shifters have already been described. Lossless mirrors have at most the effect of a phase shift. Beam splitters will be described in the next section, and then we will assemble the optical components into an interferometer.

3.7.1 Beam Splitters

Beam splitters in general We consider a *lossless beam splitter* with two input and two output radiation modes, as shown schematically in Fig. 3.7. The input photon state $|\psi\rangle$ as well as the output photon state $|\psi'\rangle$ have the form of Eq. (3.70). The assignment of the paths is carried out in accord with Fig. 3.7.

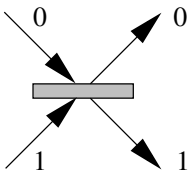


Figure 3.7: Convention for the paths of a beam splitter.

The fact that the beam splitter is lossless guarantees the conservation of the probability. When the input state is $|\psi\rangle$, then the output state $|\psi'\rangle$ is also a normalised vector and the effect of the beam splitter can thus be represented by a unitary transformation U (compare Sect. 1.1.5)

$$|\psi'\rangle = U|\psi\rangle \ . \tag{3.71}$$

We write U in the matrix form (3.36) and introduce new notations for the amplitudes and phases. ρ, τ and δ are real.

$$U = e^{i\kappa} \begin{pmatrix} \rho e^{i\delta_r} & -\tau e^{-i\delta_t} \\ \tau e^{i\delta_t} & \rho e^{-i\delta_r} \end{pmatrix} \ . \tag{3.72}$$

The action of U can be described simply when the input photon state is a vector of the computational basis:

$$|0\rangle \to U_{00}|0\rangle + U_{10}|1\rangle, \qquad |1\rangle \to U_{01}|0\rangle + U_{11}|1\rangle \ . \tag{3.73}$$

For the input state $|0\rangle$, the beam splitter causes a phase shift of $\kappa + \delta_r$ in reflection, and a multiplication by the real *reflection factor* ρ. The transmission is determined in a corresponding way by the phase shift $\kappa + \delta_t$ and the *transmission factor* τ. Analogous expressions hold for the state $|1\rangle$. The action on superpositions follows immediately.

With Eq. (3.72), we obtain as a result of conservation of probability and thus of unitarity:

$$\rho^2 + \tau^2 = 1 \; . \tag{3.74}$$

We now introduce the phase difference δ_0 between the reflected and the transmitted states for the case of the input state $|0\rangle$, and δ_1 for the input state $|1\rangle$. Eq. (3.72) leads to

$$\delta_0 = \delta_r - \delta_t \; ; \quad \delta_1 = -\delta_r + \delta_t \pm \pi \; . \tag{3.75}$$

We thus find as a further result of unitarity the relation

$$\delta_0 + \delta_1 = \pm\pi \tag{3.76}$$

which is obeyed by every beam splitter.

Special beam splitters Two types of dielectric beam splitters are often used in practice. These are firstly a beam splitter which causes a phase shift of $\frac{\pi}{2}$ on reflection in any direction and no phase shift on transmission:

$$U_1 = \begin{pmatrix} i\rho & \tau \\ \tau & i\rho \end{pmatrix} \; . \tag{3.77}$$

It corresponds to the values $\delta_r = 0$, $\delta_t = -\frac{\pi}{2}$ and $\kappa = \frac{\pi}{2}$. This beam splitter is not symmetrical in time, since $U_1^{-1} \neq U_1$. The spatially-symmetrical version of this beam splitter, in which the reflectivity and the transmissivity are the same, is given by $\rho = \tau = \frac{1}{\sqrt{2}}$ and has the effects

$$U_1|0\rangle = \frac{1}{\sqrt{2}}(i|0\rangle + |1\rangle), \tag{3.78}$$

$$U_1|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle). \tag{3.79}$$

For both input modes, a photon will be detected with the same probability $\frac{1}{2}$ in one of the output modes ($50:50$ beam splitter). We rewrite Eq. (3.77) with $\rho = \cos\theta$ and $\tau = \sin\theta$, and obtain the unitary transformation in the form

$$U_1 = (\cos\theta)\mathbb{1} - i(\sin\theta)\sigma_x = \exp(-i\theta\sigma_x) \; . \tag{3.80}$$

A global phase factor i was left off in this expression.

The other often-used type of beam splitter causes a phase jump of π on reflection from one of its sides:

$$U_2 = \begin{pmatrix} \rho & \tau \\ \tau & -\rho \end{pmatrix} \; . \tag{3.81}$$

It corresponds to the choice $\delta_r = \frac{\pi}{2}$, $\delta_t = \frac{\pi}{2}$ and $\kappa = -\frac{\pi}{2}$, and is symmetrical in time due to $U^{-1} = U$. It is, however, not spatially symmetrical, i. e. it acts differently on input photons in the modes $|0\rangle$ and $|1\rangle$. In the special case that $\rho = \tau = \frac{1}{\sqrt{2}}$, we have an *optical implementation of a Hadamard gate* H of Eq. (3.53), whose action is described by Eq. (3.54).

A universal quantum gate for qubits We write $\rho = \cos\phi$ and $\tau = \sin\phi$, and compose $\sigma_z U_2$. Explicit computation of the corresponding matrices then yields with Eq. (3.14)

$$\exp(i\phi\sigma_y) = \sigma_z U_2 \ . \tag{3.82}$$

Since σ_z can be implemented with phase shifters, and Eq. (3.37) contains (along with phase shifts $\exp(-\frac{i}{2}\lambda\sigma_z)$ and $\exp(-\frac{i}{2}\nu\sigma_z)$) only one operator of the form $\exp(i\mu\sigma_y)$, it follows that *every unitary transformation (3.71) can be implemented by a beam splitter as described by U_2 in Eq. (3.81) (Hadamard gate) along with phase shifters*.

3.7.2 Interferometer

A *Mach-Zehnder interferometer* is constructed by inserting a phase shifter between two Hadamard beam splitters, e. g. in the 1 path, whereby the $|1\rangle$ component of the photon state is multiplied by a phase factor (compare Fig. 3.8). We take a phase factor of the form $\exp(-i\alpha)$. The beam deflections by two identical ideally-reflecting mirrors have no influence on the relative phases. With the phase gate $\phi(\alpha)$ of Eq. (3.51), we can represent the interferometer symbolically by the diagram in Fig. 3.9. A set of instructions for the application of *gates* in a particular sequence is called a *circuit*. The interferometer is an example of a simple *quantum circuit*.

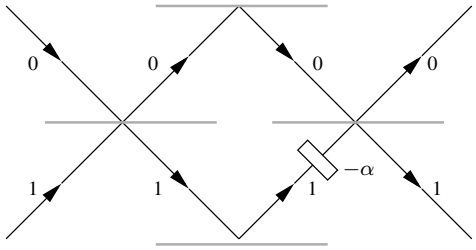


Figure 3.8: A Mach-Zehnder interferometer with phase shifter.

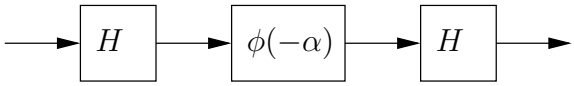


Figure 3.9: Symbolic diagram of the Mach-Zehnder interferometer.

When a photon in the state $|0\rangle$ and thus in the 0 mode enters the interferometer, it experiences in sequence the transformations

$$|0\rangle \to \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \to \frac{1}{\sqrt{2}}(|0\rangle + e^{-i\alpha}|1\rangle) \to \frac{1}{2}\left[(1+e^{-i\alpha})|0\rangle + (1-e^{-i\alpha})|1\rangle\right] \ . \tag{3.83}$$

The probability that thereafter a detector will register the photon in the state $|0\rangle$ (i. e. that a detector at the 0 output responds) is given by

$$p_0 = \frac{1}{2}(1 + \cos\alpha) \ . \tag{3.84}$$

Depending on the phase shift α, a periodically oscillating *interference pattern* results.

The interference pattern If the state $|0\rangle$ is input into the interferometer, the first Hadamard beam splitter has from Eq. (3.83) the sole function of preparing a single state vector which is symmetric in both modes. The probability that a detector in the 0 path or the 1 path behind the beam splitter emits a signal is then in each case $\frac{1}{2}$. Equation (3.83) shows that the effect of the second beam splitter is to add the complex amplitudes of the $|0\rangle$ and $|1\rangle$ components of the state vector to give the complex amplitude of the output $|0\rangle$ vector, and thus to produce interference. The interference pattern can be registered as a function of α by a detector in the 0 path on the output side of the interferometer.

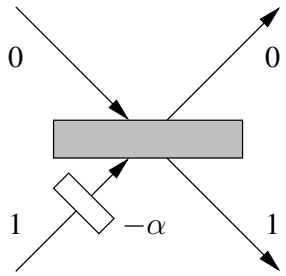


Figure 3.10: A beam splitter with a phase shifter.

The action of the second beam splitter in producing *interference* can be described in another manner, which we shall use later for other purposes. If instead of the state $|0\rangle$ a superposition of $|0\rangle$ and $|1\rangle$ is input into the interferometer, then behind the first beam splitter, there is a general state which we shall call $|\chi\rangle$. The state $|\chi\rangle$ which is input into the phase shifter and the beam splitter of the setup in Fig. 3.10 is converted by the actions of $\phi(-\alpha)$ and H, according to

$$|\chi^{\text{out}}\rangle = H(|0\rangle\langle 0| + e^{-i\alpha}|1\rangle\langle 1|)|\chi\rangle \ , \tag{3.85}$$

into the output state $|\chi^{\text{out}}\rangle$. The probability that the photon will be found in the state $|0\rangle$ is

$$p_0(\alpha) = |\langle 0|\chi^{\text{out}}\rangle|^2 \ . \tag{3.86}$$

We insert $|\chi^{\text{out}}\rangle$ and let the operators act on $|0\rangle$; then it follows that

$$p_0(\alpha) = |\langle\alpha|\chi\rangle|^2 \tag{3.87}$$

with

$$|\alpha\rangle := (|0\rangle\langle 0| + e^{i\alpha}|1\rangle\langle 1|)H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\alpha}|1\rangle) \ . \tag{3.88}$$

The measured value $p(\alpha)$ can be written directly as the mean value of the projection operator $P_{|\alpha\rangle} := |\alpha\rangle\langle\alpha|$ in the state $|\chi\rangle$ after the first beam splitter:

$$p_0(\alpha) = \langle\chi|P_{|\alpha\rangle}|\chi\rangle \ . \tag{3.89}$$

If we choose $|\alpha\rangle$ itself to be the state $|\chi\rangle$, then owing to $p_0(\alpha) = 1$, only the detector on the output 0 path registers a count. Equation (3.89) can therefore be interpreted as follows: the probability that the 0 detector will respond is equal to the probability that the input state $|\chi\rangle$ is identical to the state $|\alpha\rangle$ which yields with certainty a signal in the 0 detector.

For the input state

$$|\chi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\frac{\phi}{2}}\sin\frac{\theta}{2}|1\rangle \tag{3.90}$$

which is characterised by the parameters θ and ϕ, Eq. (3.87) can be readily evaluated, with the result

$$p_0(\alpha) = \frac{1}{2}\left[1 + \sin\theta\cos\left(\alpha - \frac{\phi}{2}\right)\right] \ . \tag{3.91}$$

On variation of α by a corresponding setting of the phase shifter, a periodic interference pattern $p_0(\alpha)$ results, with a *fringe contrast* of

$$\nu := \frac{p_{\max} - p_{\min}}{p_{\max} + p_{\min}} = \sin\theta \ . \tag{3.92}$$

By analysis of the interference pattern, the phase shift φ and the fringe contrast $\sin\theta$ can be determined. *We have thus found an interferometric procedure for the determination of the state* $|\chi\rangle$.

3.8 Locating a Bomb Without Exploding It by Using a Null Measurement*

We want to imagine a procedure which allows us in principle (at least in a certain fraction of cases) to locate extremely sensitive bombs without exploding them. Locating bombs is only a particularly spectacular example of a general application. The procedure is based on null measurements, which have other applications too, e.g. in single-photon optics.

Null measurements A *null measurement* (or negative-result measurement) is a projection measurement with a peculiarity: there exists a result for which the measuring device experiences no change, i.e. in which the indicator remains in its initial position 0. This is the *null result*. We consider a qubit. Let us assume that the null result corresponds to the state $|0\rangle$ of the quantum object. The null measurement NM is then described by:

$$\begin{aligned} \text{NM:}\ &(\text{Result } 0,\ P_0 = |0\rangle\langle 0|) \leftrightarrow \text{ measurement device unchanged} \\ &(\text{Result } 1,\ P_1 = |1\rangle\langle 1|) \leftrightarrow \text{ change in the indicator position} \ . \end{aligned}$$

*The sections marked with an asterisk * can be skipped over in a first reading.

The following example shows that the NM scheme must often be extended in the case of realistic detectors. We consider two polarisation states, $|0\rangle$ and $|1\rangle$, of a photon. A 2-level atom can be excited only by a photon with the polarisation $|1\rangle$. If it is initially in its ground state, it can absorb a $|1\rangle$ photon and thereby remove it from the radiation field. A $|0\rangle$ photon will not be absorbed and continues on its way through the apparatus. The atom in this case remains in its ground state. The ground state (excited state) of the atom then yields an unchanged (changed) indicator reading in the measurement apparatus. We shall denote such a null measurement with an additional absorption in the state $|1\rangle$ as $\mathrm{NMA}_{|1\rangle}$.

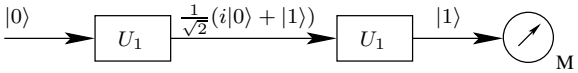


Figure 3.11: The unitary transformations U_1 of Eqs. (3.78) and (3.79) are applied to the state $|0\rangle$ of a qubit. M is a measurement in the computational basis $\{|0\rangle, |1\rangle\}$

A null measurement is detected We consider an input photon in the state $|0\rangle$ on which the unitary transformation U_1 described by Eqs. (3.78) and (3.79) acts two times (see Fig. 3.11). In a subsequent measurement M with the projection operators $|0\rangle\langle 0|$ or $|1\rangle\langle 1|$, the polarisation $|1\rangle$ is always registered – and never the polarisation $|0\rangle$.

If we insert the null measurement $\mathrm{NMA}_{|1\rangle}$ between the two unitary transformations (compare Fig. 3.12), then with equal probabilities either the photon will be absorbed or it continues on in the state $|0\rangle$. The second case is the null result. The measurement device remains unchanged. In this case, the photon is in the state $\frac{1}{\sqrt{2}}(i|0\rangle + |1\rangle)$ after the second transformation U_1. With the null measurement inserted, it is thus possible that in the measurement M, the polarisation $|0\rangle$ will be registered. Relative to the input photon, the probability for this case is $\frac{1}{4}$.

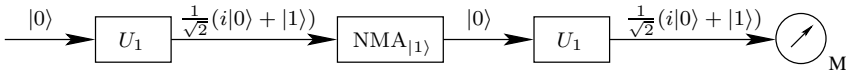


Figure 3.12: A null measurement $\mathrm{NMA}_{|1\rangle}$ with absorption in the state $|1\rangle$ of the qubit is inserted between the two unitary transformations U_1 of Fig. 3.11, which in half the cases (as shown in the figure) leads to the state $|0\rangle$.

This has a consequence of which we shall make further use. We assume that only these two cases are possible, that either the device for the measurement $\mathrm{NMA}_{|1\rangle}$ is inserted, or not. Then in a measurement M on a single photon, it can be concluded with certainty from a registration of the polarisation $|0\rangle$, that 1.) the $\mathrm{NMA}_{|1\rangle}$ device was inserted (Fig. 3.12) and 2.) that the $\mathrm{NMA}_{|1\rangle}$ device remains unchanged in its initial state. In the case of an initial state $|0\rangle$, this type of unambiguous conclusion is possible in one-fourth of all measurements.

"Detection without destruction" This leads to a surprising effect when $|0\rangle$ and $|1\rangle$ are not polarisation states, but rather as in Sect. 3.7 they refer to the spatially separated paths in

an interferometer which is constructed from mirrors and beam splitters of the type U_1 (phase shifts of $\frac{\pi}{2}$ on reflection.) A bomb in the optical path 1 (cf. Fig. 3.13), which is so sensitive that it will certainly explode on interacting with only a single photon, is equivalent to an $\mathrm{NMA}_{|1\rangle}$ device for the null measurement with absorption. The immediate analogy to the situation of polarised photons described above demonstrates this.

We make use of this null measurement to locate bombs. To this end, we place the interferometer in such a way that the optical path 1 leads through a region in which we presume that a bomb may be present. If the detector D_0 registers a signal, then we can conclude with certainty, without having triggered the bomb, that there is in fact a bomb in this region. This applies in one-fourth of the measurements. In half of all the measurements, this procedure results in the triggering of the bomb which is present. In the remaining one-fourth of the measurements, the detector D_1 responds. Then we can reach no conclusion.

The fact that in individual cases a signal from a detector (here the detector D_0) allows the certain conclusion that a bomb exists without its being triggered is a quantum effect. It is based on the use of a single photon, on the superposition by the beam splitter, and on the fact that the presence of the bomb in one of the optical paths gives rise to a null measurement[2]. There is a result in which the measurement device remains unaffected. At the same time, the object (photon) is not destroyed, but rather is transferred to a new state. It can thereby pass on the information that a measurement has taken place.

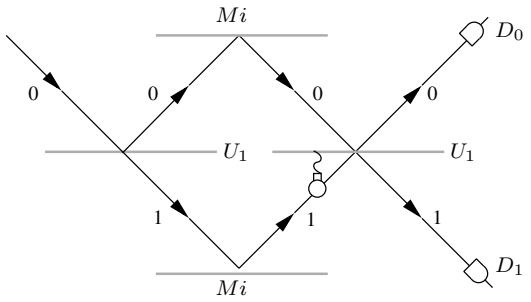


Figure 3.13: The quantum circuit of Fig. 3.12 is implemented in terms of an interferometer with a bomb in optical path 1. Both the U_1 are 50:50 beam splitters with a phase shift of $\frac{\pi}{2}$ on reflection. The two identical mirrors are denoted by Mi. D_0 and D_1 are detectors of single photons.

[2]The bomb would be triggered by a single photon. If the photon is registered in detector D_0, it cannot have collided with the bomb. Are we thus dealing with an *interaction-free measurement*? Those who have already read Chap. 15 will know that the measurement begins with the entanglement (see Eq. (15.24)), which in the case of the null measurement takes on the form

$$(c_0|0\rangle + c_1|1\rangle)|i^M\rangle \xrightarrow{NM} c_0|0\rangle|i^M\rangle + c_1|1\rangle|1^M\rangle \; . \tag{3.93}$$

Here, $|i^M\rangle$ is the state of the measurement device, which remains unchanged during the measurement (null result). The entanglement of Eq. (3.93) requires an interaction between the quantum system and the measurement device. The subsequent decoherence is an interaction between the measurement device and its environment. In the example with the bomb as measurement device, we presumed that these interactions do not cause the bomb to explode. The quantum-mechanical treatment of a null measurement shows that it is not interaction free. Nevertheless, the null result corresponding to an unaffected detector can occur with a probability of $|c_0|^2$.

3.9 Complementary Topics and Further Reading

- Systems with spins greater than $\frac{1}{2}$ are investigated in detail in [Zei 81], [PSM 87], [CST 89], [MW 95] with respect to their role as subsystems in composite quantum systems.
- Interactions between light and a 2-level atom: [MW 95].
- Review articles on quantum gates: [Bra 02].
- Generalised lossless beam splitters: [CST 89], [MW 95].
- The bomb search was suggested in [EV 93]. An overview of the literature on the theory and on experiments can be found in [Vai 03]. There, the justification of the term "interaction-free measurement" is discussed.
- By making use of the quantum Zeno effect, the efficiency of the bomb search can be increased [KWM 99].
- Employing states which contain more than one photon allows one to carry out more complex null measurements and to make practical use of them [KSN 06].
- A technical application of "interaction-free" measurements could be "interaction-free" imaging, in which only a small number of photons need interact with the object being imaged [WMN 98]. The amount of radiation damage can thus be reduced.

3.10 Problems for Chapter 3

Prob. 3.1 [for Sects. 3.1 and 3.3]: Derive equations (3.11), (3.17), (3.36), (3.37), (3.42), (3.55), and (3.56).

Prob. 3.2 [for Sect. 3.1]: Determine the components of the operator $\sigma_i \sigma_j$ in the operator basis of the σ operators.

Prob. 3.3 [for Sect. 3.4]: Compute the representations of the operators in Tab. 3.1, so far as this was not already done in the text.

Prob. 3.4 [for Sect. 3.4]: Implement for a single photon a $\sqrt{\text{NOT}}$ gate using a beam splitter. Construct a NOT gate from two beam splitters.

Prob. 3.5 [for Sect. 3.7.2]: In the 0 path of a Mach-Zehnder interferometer, a phase shifter with the effect $e^{i\alpha}$ has been inserted, and in the 1 path is another with the effect $e^{-i\alpha}$. Show that the resulting transformation has the form

$$U = \begin{pmatrix} \cos\alpha & i\sin\alpha \\ i\sin\alpha & \cos\alpha \end{pmatrix} . \tag{3.94}$$

By insertion of additional phase shifters in the input and output modes, an arbitrary unitary transformation can be implemented.

Begin with the beam splitter described by the transformation U_2 in Eq. (3.81) and show that by insertion of phase shifters, the beam splitter giving the transformation U_1 in Eq. (3.77) can be constructed.

© 2007 WILEY-VCH Verlag GmbH & Co.

4 Mixed States and the Density Operator

We have already met up with mixed states in the introductory Sect. 2.1.1, but there, we restricted our formulation of the postulates to pure states. These pure states were described by vectors in the Hilbert space. We will now first introduce a different representation of the pure states, which will then lead us directly to an approach for the description of statistical mixtures (blends) and general mixtures by using density operators. Density operators describe the general quantum state. The postulates from Sect. 2.1.2 will be generalised in the following.

4.1 Density Operators for a Given Ensemble (Statistical Mixture)

4.1.1 Pure States

To a pure state, which we have up to now described by the normalised state vector $|\psi\rangle$, we can also unambiguously assign the operator

$$\rho := |\psi\rangle\langle\psi| \ . \tag{4.1}$$

It is called the *density operator of a pure state* or often also the *density matrix*. The following properties of the density operator can be read off directly:

(i) ρ is positive: $\langle\varphi|\rho|\varphi\rangle \geq 0 \ , \forall \ |\varphi\rangle \in \mathcal{H}_d$ (and thus Hermitian, $\rho^\dagger = \rho$)

(ii) $tr[\rho] = 1$

(iii) $\rho^2 = \rho$.

Property (ii) is a result of the normalisation of $|\psi\rangle$. Conversely, the three properties together guarantee that the spectral decomposition of ρ has the form (4.1), and thus that ρ uniquely determines the vector $|\psi\rangle$ up to a phase. Making use of the spectral decomposition, one can also show that when (i) and (ii) hold for an operator ρ, then property (iii) is equivalent to

$$\text{(iii}^*\text{)} \ \text{tr}[\rho^2] = 1 \ . \tag{4.2}$$

A selective measurement of the observables A with the result a_n and the associated projection operator P_n from Eq. (2.3) converts ρ, according to Eq. (2.2), into the density operator ρ'_n

$$\rho \rightarrow \rho'_n = \frac{1}{p(a_n)} P_n \rho P_n \ . \tag{4.3}$$

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

Here, corresponding to Eq. (2.5),

$$p(a_n) = \mathrm{tr}[P_n \rho] \tag{4.4}$$

is the probability of obtaining the value a_n in the measurement. It can be useful to write the resulting density operator in its *non-normalised form* $\tilde{\rho}'_n := |\tilde{\psi}'_n\rangle\langle\tilde{\psi}'_n|$, with $|\tilde{\psi}'_n\rangle$ from Eq. (2.2). In this form, the trace is not necessarily equal to one. As with the state vector, we characterise this by a tilde:

$$\tilde{\rho}'_n = P_n \rho P_n \ . \tag{4.5}$$

The probability $p(a_n)$ is then equal to the trace of the non-normalised resulting density operator after the selective measurement:

$$p(a_n) = \mathrm{tr}[\tilde{\rho}'_n] \ . \tag{4.6}$$

We obtain the expectation value $\langle A \rangle$ of the observable A by inserting the identity operator generated by the ONB $\{|u_i\rangle\}$,

$$\begin{aligned} \langle A \rangle &= \sum_{j,k=1}^{d} \langle\psi|u_j\rangle\langle u_j|A|u_k\rangle\langle u_k|\psi\rangle \\ &= \sum_{j,k} \langle u_k|\rho|u_j\rangle\langle u_j|A|u_k\rangle \\ &= \sum_{k} \langle u_k|\rho A|u_k\rangle \\ &= \mathrm{tr}[\rho A] \ . \end{aligned} \tag{4.7}$$

We still have the task of rewriting the unitary dynamics in the Schrödinger representation for ρ. With Eq. (2.10), we find

$$\rho(t) = U(t,t_0)\rho(t_0)U^{-1}(t,t_0) \ , \tag{4.8}$$

and we obtain the *von Neumann equation*:

$$\begin{aligned} i\hbar\dot{\rho}(t) &= i\hbar\dot{U}\rho(t_0)U^{-1} + i\hbar U\rho(t_0)\dot{U}^{-1} \\ &= HU\rho(t_0)U^{-1} - U\rho(t_0)U^{-1}H \\ &= [H, \rho(t)] \ . \end{aligned} \tag{4.9}$$

With the Liouville operator $\mathcal{L}$ as in Eq. (1.87), it can also be written in the form

$$i\hbar\dot{\rho} = \mathcal{L}(\rho) \ . \tag{4.10}$$

To summarise, we have seen that the physics of quantum systems in pure states, which can be described by the mathematical object "state vector $|\psi\rangle$", can be equally well specified by the density operator ρ as in Eq. (4.1). In this sense, one can say with the same operational significance as for $|\psi\rangle$: the system is in the *state* ρ. In contrast to the vector formulation, we can however apply the state formulation in terms of the density operator directly to a more general class of quantum states, the statistical mixtures. We shall describe this in detail in the following section.

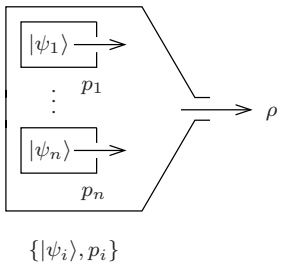


Figure 4.1: The statistical mixture (blend) with ensemble $\{|\psi_i\rangle, p_i\}$ results if only one of the preparation apparatus for the states $|\psi_i\rangle$ acts at any given time, with the corresponding probability p_i.

4.1.2 The Physics of Statistical Mixtures (Blends)

Preparation We consider the following experimental situation: for isolated quantum systems of the same type, an arbitrary but finite number of different preparation procedures are available, which are enumerated by the index i ($i = 1, \ldots, N$), and which convert it correspondingly into the pure states $|\psi_i\rangle$. These states need be neither orthogonal nor linearly independent. N can be greater than the dimension d. We proceed to a new type of preparation procedure which consists in applying for each quantum system one of the original procedures with a particular *classical probability* p_i, where

$$\sum_{i=1}^{N} p_i = 1 \, . \tag{4.11}$$

Therefore, for the preparation of the single quantum system under consideration, precisely one of the preparation apparatus is switched on in a random manner (compare Fig. 4.1). *The corresponding state $|\psi_i\rangle$ is then in fact realised for the respective system.* In this process, it is guaranteed that the $i\ th$ apparatus acts with the probability p_i. One says that through "mixing" of the pure states, the *ensemble* $\{|\psi_i\rangle, p_i\}$ is generated. We have already encountered an example of this in Sect. 2.1.1. With knowledge of the states $|\psi_i\rangle$ and of the preparation probabilities p_i, a definite prediction of the probability of occurrence of measured results is again possible. *According to the concept which we are using, this generalised preparation procedure thus defines a quantum state*. It is called a *statistical mixture* or *blend* and can be a pure state as a special case. By the use of the term "statistical", we emphasize that an ensemble has been prepared. In the literature, one can also find the term *proper mixture*. In Chap. 7, we shall meet up with another type of mixture.

The density operator of statistical mixtures We show that the equations for the physical expressions of Sect. 4.1.1 can be applied to statistical mixtures when they are described by the *density operator*

$$\rho := \sum_{i=1}^{N} p_i |\psi_i\rangle\langle\psi_i| = \sum_{i=1}^{N} p_i \rho_i \tag{4.12}$$

with $\rho_i := |\psi_i\rangle\langle\psi_i|$. To this end, we reduce statements about the mixture to statements about the ensemble states ρ_i.

In a measurement of the observable A, the value a_n will be obtained with the probability

$$p(a_n) = \sum_i p(a_n|i)p_i \ . \tag{4.13}$$

Here, $p(a_n|i)$ is the *conditional probability* that the value a_n will be obtained when the state ρ_i is present. With (4.4), we find

$$p(a_n|i) = \mathrm{tr}[P_n\rho_i] \ . \tag{4.14}$$

This leads with Eq. (4.12) and the rules for computation of the trace to

$$p(a_n) = \mathrm{tr}[P_n\rho] \ . \tag{4.15}$$

Equation (4.4) can thus be applied to the case of mixtures. Equation (4.7) can also be applied to the case of statistical mixtures with the argument that expectation values for the individual states contribute with the probabilities p_i to the expectation value for the mixture:

$$\langle A\rangle = \sum_i p_i \mathrm{tr}[A\rho_i] = \mathrm{tr}[A\rho] \ . \tag{4.16}$$

As shown by Eqs. (4.13) and (4.16), in the computation of the probabilities and expectation values, products of classical and quantum-mechanical probabilities occur.

Selective and non-selective measurements We still need to derive the equations for the two forms of dynamic behaviour. We begin with the measurement dynamics. For each of the individual states ρ_i of the ensemble we obtain for the result of a *selective measurement* giving the value a_n the non-normalised state

$$\rho_i \rightarrow \tilde{\rho}'_{i,n} = P_n\rho_i P_n \tag{4.17}$$

(compare Eq. (4.5)). We explicitly allow degeneracy of the measured values. For the density operator ρ of Eq. (4.12), it follows from Eq. (4.17) that

$$\rho \rightarrow \tilde{\rho}'_n = \sum_i p_i \tilde{\rho}'_{i,n} = P_n\rho P_n \ . \tag{4.18}$$

This agrees with the relation (4.5). With Eq. (4.15), we find for the normalised density operator following the selective measurement, in agreement with Eq. (4.3),

$$\rho \rightarrow \rho'_n = \frac{1}{p(a_n)} P_n\rho P_n \ , \quad \mathrm{tr}[\rho'_n] = 1 \ . \tag{4.19}$$

A derivation of this equation starting from the determining quantity, i.e. from the measurement result a_n, can also be carried out by applying Bayes' theorem[1] (see Sect. 1.3.2). The

[1] Assume that a measurement of the observable A leads to the result a_n. If the state ρ_i was present before the measurement, then after the measurement, the normalised state

$$\rho'_{i,n} = \frac{1}{p(a_n|i)} P_n\rho_i P_n \tag{4.20}$$

probability $p(a_n)$ of observing the measured value a_n can, according to Eqs. (4.15) and (4.18) and in agreement with Eq. (4.6), also be written in the form

$$p(a_n) = \mathrm{tr}[\tilde{\rho}'_n] \; . \tag{4.23}$$

Finally, we mention also the special case that no degeneracy is present for the measured value a_n. Then we have $P_n = |a_n\rangle\langle a_n|$ with the eigenvector $|a_n\rangle$ belonging to the eigenvalue a_n. Accordingly, Eq. (4.19) becomes

$$\rho \to \rho'_n = |a_n\rangle\langle a_n| \; . \tag{4.24}$$

When the result of a measurement is not degenerate, the corresponding selective measurement on a statistical mixture prepares a pure state.. This is plausible, since in this case each of the state vectors $|\psi_i\rangle$ of the ensemble is converted by the measurement process to $|a_n\rangle$.

In a *non-selective measurement* of the observable A, one often repeats the measurement on quantum systems in the same state without sorting out those states ρ'_n which belong to a particular result a_n as in a selective measurement. The resulting state ρ' in this case is again a statistical mixture which is composed additively of the states ρ'_n that are produced by the measurement with the probabilities $p(a_n)$:

$$\rho \xrightarrow{n.s.} \rho'_{n.s.} = \sum_n p(a_n) \frac{P_n \rho P_n}{\mathrm{tr}[\rho P_n]} = \sum_n P_n \rho P_n, \qquad \mathrm{tr}[\rho'_{n.s.}] = 1 \; . \tag{4.25}$$

We made use of Eq. (4.15) in deriving this result.

Unitary dynamics We now turn to the unitary dynamics. Under their influence, the classical probabilities p_i of the preparation procedure do not change. Equation (4.8) can thus be directly applied to statistical mixtures

$$\rho(t) = U(t, t_0)\rho(t_0)U^{-1}(t, t_0) \tag{4.26}$$

and we find the von Neumann equation (4.9) or (4.10).

(cf. Eqs. (4.14) and (4.23)) is found. What is the probability $p(i|)a_n)$ that ρ_i was present before the measurement, if a_n resulted from the measurement? According to Bayes' theorem, which was treated in more detail in Sect. 1.3.2, we find (by rearranging with the help of Eq. (4.13))

$$p(i|a_n) = \frac{p(a_n|i)p_i}{\sum_j p(a_n|j)p_j} = \frac{p(a_n|i)p_i}{p(a_n)} \; , \tag{4.21}$$

where $\sum_i p(i|a_n) = 1$. The quantity $p(i|a_n)$ is thus in general not the same as p_i. Overall, it follows that the measurement resulting in a_n belongs to the transition

$$\rho \to \rho'_n = \sum_i p(i|a_n)\rho'_{i,n} \; . \tag{4.22}$$

Substitution of $\rho'_{i,n}$ leads directly to equation (4.19).

Where do statistical mixtures occur? The preparation procedure described at the beginning of this chapter may appear somewhat artificial. However, as we have seen, for isolated quantum systems, statistical mixtures indeed occur in a very natural manner. All non-selective measurements on pure states and on statistical mixtures lead to statistical mixtures and not to pure states.

Even selective measurements on statistical mixtures in general lead to statistical mixtures when the resulting measured value is a degenerate eigenvalue of the observable operators. Degeneracy can lead to mixing. With ρ from Eq. (4.12), we find for ρ'_n in Eq. (4.19)

$$\rho'_n = \frac{1}{p(a_n)} \sum_i p_i P_n |\psi_i\rangle\langle\psi_i| P_n \ . \tag{4.27}$$

Projection operators P_n for degenerate eigenvalues project onto subspaces (cf. Eq. (2.3)), therefore the $P_n|\psi_i\rangle$ are in general not all the same. The resulting density operator is a sum of density operators for pure states and can thus not itself be the density operator of a pure state. We discuss this point below in more detail in connection with convex combinations. In the presence of degeneracy, the mixture is not completely unmixed by a selective measurement.

Finally, we mention another measurement situation which leads to statistical mixtures. Assume the measured values obtained by a measurement apparatus to be non-degenerate. However, there is an inaccuracy in the indication of the apparatus, so that all the values within the interval $[a_n, a_{m>n}]$ are indicated by the apparatus as a single result a_n. A selective measurement yielding the apparent value a_n then does not prepare a pure state, but rather a statistical mixture.

We have up to now considered only isolated systems. In Sect. 7.3, we shall see that the subsystems of entangled systems can likewise be described by density operators. This leads to an important extension of the concept of a "mixture" to include non-statistical mixtures which do not correspond to any well-determined ensemble of individually prepared states.

4.1.3 Definition and Properties of the Generalised Density Operator

Definition We now must determine which of the properties (i) to (iii$^{(*)}$) that were derived for pure states in Sect. 4.1.1 also hold for the density operator of a statistical mixture. Making use of Eqs. (4.11) and (4.12), we confirm (i) and (ii) immediately. For the discussion of (iii), we consider the spectral decomposition of ρ,

$$\rho = \sum_{j=1}^{d} \lambda_j |j\rangle\langle j| \ . \tag{4.28}$$

$\{|j\rangle\}$ is an ONB and as a result of (i) and (ii), we have $\lambda_j = \lambda_j^*$, $\lambda_j \geq 0$ and $\sum_j \lambda_j = 1$. We thus obtain

$$0 \leq \lambda_j \leq 1 \tag{4.29}$$

and thereby

$$\mathrm{tr}[\rho^2] = \sum_j \lambda_j^2 \leq 1 \ . \tag{4.30}$$

The equals sign in Eq. (4.30) unambiguously characterises the occurrence of a pure state. The inequality holds only for a true mixture (never for a pure state).

We discard the experimental implementations considered up to now and call an operator quite generally a *density operator* if it fulfills the conditions

(i) ρ is positive (and thus Hermitian, $\rho^\dagger = \rho$)

(ii) $\mathrm{tr}[\rho] = 1$.

The inequality (4.30) is a consequence of this.

Degree of mixtureness From Eq. (4.30), we find for the smallest value of $\mathrm{tr}[\rho^2]$ the quotient $\frac{1}{d}$, where d is the dimension of the Hilbert space. The value $\lambda_j = \frac{1}{d}$ is adopted and belongs to

$$\rho = \frac{1}{d}\mathbb{1} \; . \tag{4.31}$$

This completely structureless density operator is called the *maximally mixed density operator*.

As described above, statistical mixtures are produced operationally in experiments by the "mixing" of states. One can introduce the parameter

$$\Xi := 1 - \mathrm{tr}[\rho^2] \tag{4.32}$$

to describe the *degree of mixtureness* which varies between that of a pure state, $\Xi = 0$, and that of maximal mixtureness, $\Xi = 1 - \frac{1}{d}$:

$$0 \leq \Xi \leq 1 - \frac{1}{d} \; . \tag{4.33}$$

Convex combinations We make note of a simple consequence of the definition of the density operator. When ρ_l with $l = 1, \ldots, k$ are density operators and r_l are positive numbers with $\sum_l r_l = 1$, then the *convex sum*

$$\rho = \sum_{l=1}^{k} r_l \rho_l \tag{4.34}$$

is again a density operator.

We show that the density operator ρ of a pure state is characterised not only by $\mathrm{tr}[\rho^2] = 1$, but also in another way. *In contrast to all other density operators, the density operator of a pure state cannot be decomposed into a convex sum.* To prove this statement, we attempt to decompose the density operator

$$\rho = |\psi\rangle\langle\psi| \tag{4.35}$$

according to

$$\rho = \lambda\rho_1 + (1-\lambda)\rho_2, \qquad 0 < \lambda < 1 \, , \quad \rho_1 \neq \rho_2 \; . \tag{4.36}$$

For a vector $|\chi\rangle$ orthogonal to $|\psi\rangle$, we obtain

$$\langle\chi|\rho|\chi\rangle = 0 = \lambda\langle\chi|\rho_1|\chi\rangle + (1-\lambda)\langle\chi|\rho_2|\chi\rangle\,. \tag{4.37}$$

Since λ and $1-\lambda$ are positive and the operators ρ_1 and ρ_2 are positive operators, it follows that

$$\langle\chi|\rho_1|\chi\rangle = \langle\chi|\rho_2|\chi\rangle = 0\,. \tag{4.38}$$

We complete $|\psi\rangle$ to an ONB by including additional vectors. For these vectors, Eq. (4.38) applies in each case. We take the matrix elements of ρ_1 and ρ_2 in this basis and make use of $\mathrm{tr}[\rho_1] = \mathrm{tr}[\rho_2] = 1$. Then we find as the only non-vanishing matrix elements

$$\langle\psi|\rho_1|\psi\rangle = \langle\psi|\rho_2|\psi\rangle = 1\,. \tag{4.39}$$

From this, we have

$$\rho = \rho_1 = \rho_2\,. \tag{4.40}$$

The decomposition (4.36) is thus not possible. *Mixing of pure states (or of mixtures) can never lead to another pure state.* This characterises pure states both mathematically as well as in principle also operationally. We made use of this property in the definition of a pure state when we formulated the postulates in Sect. 2.1.2.

4.1.4 Incoherent Superpositions of Pure States

During a unitary dynamic evolution, as a consequence of Eq. (4.26), we have conservation of the positivity and of the trace of ρ:

$$\mathrm{tr}[\rho(t)] = \mathrm{tr}[\rho(t_0)] \tag{4.41}$$

as well as of the trace of ρ^2

$$\mathrm{tr}[\rho^2(t)] = \mathrm{tr}[\rho^2(t_0)]\,. \tag{4.42}$$

Therefore, by unitary dynamics, neither can a pure state be transformed into a mixture, nor vice versa.

On the other hand, as a result of a selective measurement with a non-degenerate measured value, a true mixture can indeed be transformed into a pure state, as we have seen in Eq. (4.24). Is the reverse process, which we call *decoherence*, possible as a result of a measurement, i.e. the second form of dynamics described by the postulates? We consider this question more thoroughly in connection with the theory of measurements in Chap. 15. Here, we initially wish to demonstrate why it makes sense to use the terms coherence and incoherence.

We can read off from Eq. (4.16) that when we take an expectation value $\langle A\rangle$, the averaging occurs over the expectation values $\mathrm{tr}[A\rho_i]$ and not, as in a superposition, over the states themselves. In the case of a statistical mixture, we are dealing in this sense with an *incoherent superposition* of pure states. The states involved do not interfere with each other. Their

relative phase cannot be experimentally determined. We can immediately understand this in light of the preparation procedure given in Sect. 4.1.2. This statement can be visualised operationally in an interference experiment. The situation of the two-slit experiment provides a further example (see Sect. 2.1).

In Sect. 3.7.2, we saw how by using a beam splitter, the 0 component and the 1 component of a photon state $|\psi^{\text{in}}\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\frac{\phi}{2}}\sin\frac{\theta}{2}|1\rangle$ can be coherently superposed. The phase shifter in Fig. 3.10 produced an interference pattern which depends on the phase α. In the special cases $|\psi^{\text{in}}\rangle = |0\rangle$ and $|\psi^{\text{in}}\rangle = |1\rangle$, no interference pattern with fringe contrast is obtained. For an input mixture ρ^{in}, according to Eq. (3.86), the probability that the 0 detector registers a signal as a function of the phase shift α is given by

$$p(\alpha) = \text{tr}[P_{|\alpha\rangle}\rho^{\text{in}}] \,, \tag{4.43}$$

where $P_{|\alpha\rangle} = |\alpha\rangle\langle\alpha|$ from Eq. (3.88). For $\rho^{\text{in}} = |0\rangle\langle 0|$ and $\rho^{\text{in}} = |1\rangle\langle 1|$, the result is $p(\alpha) = \frac{1}{2}$, i.e. there is no interference pattern.

If we do not superpose the states $|0\rangle$ and $|1\rangle$ as in the case of $|\psi^{\text{in}}\rangle$, but instead mix them:

$$\rho^{\text{in}} = \lambda_0|0\rangle\langle 0| + \lambda_1|1\rangle\langle 1| \tag{4.44}$$

($\lambda_{0,1} \geq 0$, $\lambda_0 + \lambda_1 = 1$), then we obtain for the signal probability once again a value which is independent of α,

$$p_0(\alpha) = \frac{1}{2}(\lambda_0 + \lambda_1) = \frac{1}{2} \,. \tag{4.45}$$

The statistical mixture ρ^{in} is an incoherent superposition for which in the interferometer with the paths $|0\rangle$ and $|1\rangle$, no interference pattern depending on α occurs. This is plausible if one imagines that for a statistical mixture, objects in the states $|0\rangle$ and $|1\rangle$ enter one after another.

Remarkably, we find the same result even for statistical mixtures with an ensemble $\{|\psi_i\rangle, p_i\}$ which is not the same as the ensemble $\{|0\rangle, |1\rangle, \lambda_0, \lambda_1\}$:

$$\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|, \tag{4.46}$$

provided that the associated density operator ρ is mathematically identical with the density operator ρ^{in}. With Eq. (4.43), it follows again that $p(\alpha) = \frac{1}{2}$. In these cases also, the interference pattern generated from all the measured points exhibits no fringe contrast, although possibly the ensemble states $|\psi_i\rangle$ individually might very well lead to interference with fringe contrast. As shown by Eqs. (4.43) and (4.46), the resulting interference pattern $p(\alpha)$ is produced by additive – and thus incoherent – superposition, $p(\alpha) = \sum_i p_i(\alpha)$, of the individual interference patterns of the states $|\psi_i\rangle$

$$p_i(\alpha) = p_i\text{tr}[P_{|}\alpha\rangle|\psi_i\rangle\langle\psi_i|] \tag{4.47}$$

weighted by the ensemble probabilities p_i. We shall return to this point in connection with the quantum eraser in Sect. 8.7.

4.2 The Generalised Quantum State

We recall the definition of the state of a quantum system as given in Section 2.1.2. The state is that mathematical object which permits us to compute the probabilities of the results of all possible measurements on the system. It is associated with a preparation procedure.

The measurement postulate *Density operators, i.e. all positive (and thus Hermitian) operators ρ, which fulfill the condition tr$[\rho] = 1$, are clearly such mathematical objects, if one postulates that the probabilities are given by Eq. (4.15). The measurement process transforms the states into the states ρ'_n of Eq. (4.19) or $\rho'_{n.s.}$ of Eq. (4.25).* This generalises postulate 2 from Sect. 2.1.2. *It can be shown that aside from the density operators, there are no other mathematical objects which fulfill the requirements of a quantum state.* According to this important *Gleason's theorem* (see Sect. 4.5), we have already arrived at the most general description of quantum systems.

Up to now, we have considered only statistical mixtures (i. e. blends) as quantum states. In Chap. 7, we will deal with other realisations of states based on the reduced density operators for the subsystems of composite systems. They differ from statistical mixtures in terms of their preparation. The postulate which generalises postulate 1 of Sect. 2.1 can then in general terms be stated as: *quantum states are represented by density operators.* They are called *mixtures*. Statistical mixtures are a special physical case which is distinguished by a particular preparation procedure.

4.3 Different Ensemble Decompositions of a Density Operator and the Ignorance Interpretation

We begin with a very simple observation. Let the decomposition of two qubit states $|a\rangle$ and $|b\rangle$ (which describe e. g. spin-$\frac{1}{2}$ states) in the computational basis be

$$\begin{aligned} |a\rangle &= \sqrt{\frac{2}{3}}|0\rangle + \sqrt{\frac{1}{3}}|1\rangle \\ |b\rangle &= \sqrt{\frac{2}{3}}|0\rangle - \sqrt{\frac{1}{3}}|1\rangle \,. \end{aligned} \tag{4.48}$$

Then the density operator ρ, which belongs to the ensemble of the states $|a\rangle$ and $|b\rangle$ with the probabilities $p_a = p_b = \frac{1}{2}$, can be written in a simple form:

$$\rho = \frac{1}{2}|a\rangle\langle a| + \frac{1}{2}|b\rangle\langle b| = \frac{2}{3}|0\rangle\langle 0| + \frac{1}{3}|1\rangle\langle 1| \,. \tag{4.49}$$

It is thus at the same time the density operator for the ensemble with the states $|0\rangle$ and $|1\rangle$ (e. g. spin polarisations in the z direction) and the probabilities $p_0 = \frac{2}{3}$ and $p_1 = \frac{1}{3}$.

We give a second physical example. In a particular experimental procedure, horizontally- and vertically-polarised photons are produced with equal probabilities. The associated ensemble $\{|H\rangle, |V\rangle, p_H = p_V = \frac{1}{2}\}$ is described by the density operator

$$\rho = \frac{1}{2}\mathbb{1} \,. \tag{4.50}$$

In a completely different setup, right-hand circular- and left-hand circular-polarised photons are produced in equal numbers. The ensemble $\{|R\rangle, |L\rangle, p_R = p_L = \frac{1}{2}\}$ has the same density operator ρ. The photons are therefore in the same state. There is no experiment which can be carried out on the photons to distinguish by which of the two experimental procedures they were prepared. Knowledge of the density operator ρ does not allow us in this case to conclude unambiguously which ensemble is present. We shall prove that this holds for every density operator and shall show at the same time how the different ensembles arise from one another. In the following, we presume that the density operator does not describe a pure state.

Ensemble decompositions We consider the density operator

$$\rho = \sum_a p_a |\psi_a\rangle\langle\psi_a| = \sum_a |\tilde{\psi}_a\rangle\langle\tilde{\psi}_a|, \quad |\tilde{\psi}_a\rangle := \sqrt{p_a}|\psi_a\rangle, \tag{4.51}$$

which we formulate with non-normalised vectors, denoted by a tilde. Again, the vectors $|\psi_a\rangle$ need not be orthogonal nor linearly independent. The fact that a density operator ρ can be considered to be the density operator of the ensembles $\{|\psi_a\rangle, p_a\}$ and can therefore be written as in Eq. (4.51) is called an *ensemble decomposition* of ρ. We shall assume that there is a further ensemble decomposition of ρ:

$$\rho = \sum_i |\tilde{\varphi}_i\rangle\langle\tilde{\varphi}_i| \ . \tag{4.52}$$

Furthermore, we always have the spectral decomposition of ρ,

$$\rho = \sum_{n=1}^{d} \lambda_n |n\rangle\langle n| = \sum_{n=1}^{d} |\tilde{n}\rangle\langle\tilde{n}| \ , \tag{4.53}$$

with the ONB $\{|n\rangle\}$, which likewise represents an ensemble decomposition of ρ. The ranges of the indices of the types $a, b, \ldots$ and $i, j, \ldots$ as well as $n, m, \ldots$ need not be the same.

We limit ourselves to eigenvalues $\lambda_n \neq 0$ and limit the range of $n, m, \ldots$ correspondingly. Then the associated $|n\rangle$ do not necessarily form a basis of the whole Hilbert space $\mathcal{H}$. We assume that they span the subspace $\mathcal{H}'$. For $|\chi\rangle$ from the orthogonal complement of $\mathcal{H}'$, we then find:

$$0 = \langle\chi|\rho|\chi\rangle = \sum_a \langle\chi|\tilde{\psi}_a\rangle\langle\tilde{\psi}_a|\chi\rangle = \sum_a |\langle\chi|\tilde{\psi}_a\rangle|^2 \tag{4.54}$$

and therefore

$$\langle\chi|\tilde{\psi}_a\rangle = 0 \tag{4.55}$$

for all $|\chi\rangle$ and all indices a. The corresponding conclusion holds for the vectors $|\tilde{\varphi}_i\rangle$. Then all the $|\tilde{\psi}_a\rangle$ and all the $|\tilde{\varphi}_i\rangle$ lie in the subspace $\mathcal{H}'$ and we can expand them in terms of the basis $\{|n\rangle\}$ of $\mathcal{H}'$:

$$|\tilde{\psi}_a\rangle = \sum_n c_{an}|\tilde{n}\rangle, \quad |\tilde{\varphi}_i\rangle = \sum_n d_{\text{in}}|\tilde{n}\rangle \ . \tag{4.56}$$

Inserting into Eq. (4.51) and taking Eq. (4.53) into account leads to

$$\rho = \sum_{a,n,m} c_{an} c_{am}^* |\tilde{n}\rangle\langle\tilde{m}| = \sum_n |\tilde{n}\rangle\langle\tilde{n}| \tag{4.57}$$

and hence to

$$\sum_a c_{an} c_{am}^* = \delta_{nm} \,, \quad \sum_i d_{\text{in}} d_{im}^* = \delta_{nm} \,. \tag{4.58}$$

The relations that follow from Eq. (4.52) are also shown in (4.58). When the ranges of the indices are the same, the Eqs. (4.58) imply that the matrices c_{an} and d_{in} are unitary.

With Eqs. (4.56) and (4.58), we find

$$\sum_a c_{am}^* |\tilde{\psi}_a\rangle = \sum_{a,n} c_{am}^* c_{an} |\tilde{n}\rangle = |\tilde{m}\rangle \,. \tag{4.59}$$

Since Eq. (4.59) holds for every basis vector of $\mathcal{H}'$, the number of vectors $|\psi_a\rangle$ and $|\varphi_i\rangle$ cannot be less than the dimension of $\mathcal{H}'$. Finally, we insert Eq. (4.59) into Eq. (4.56) and obtain

$$|\tilde{\varphi}_i\rangle = \sum_{n,a} d_{\text{in}} c_{an}^* |\tilde{\psi}_a\rangle \,. \tag{4.60}$$

A corresponding decomposition can be obtained by applying Eq. (4.58) for $|\tilde{\psi}_a\rangle$.

We have shown that: *For two ensemble decompositions $\{|\tilde{\psi}_a\rangle\}$ and $\{|\tilde{\varphi}_i\rangle\}$ of a density operator, the vectors of the one decomposition can be written as a linear combination of the vectors of the other decomposition according to Eq. (4.60), whereby Eq. (4.58) is obeyed. Conversely, one can readily show that vectors $|\tilde{\varphi}_i\rangle$ and $|\tilde{\psi}_a\rangle$, which are related as in Eq. (4.60), each represent an ensemble decomposition of the same density operator when their matrices fulfill the conditions (4.58). Evidently there are arbitrarily many such matrices and therefore arbitrarily many ensemble decompositions of a density operator.*

The example of Eq. (4.49) and the proof have made it clear that the ambiguity of the ensemble decomposition has an underlying cause which is typical of quantum mechanics: a state vector can be written in infinitely many different ways as linear combinations of other state vectors. *There is no analogy for classical states.*

The ignorance interpretation We have seen that a density operator permits mathematically different ensemble decompositions. This should not lead to confusion in connection with the question as to what is really present in the physical world. When quantum systems are prepared in the state ρ as a statistical mixture (or a blend) with the ensemble $\{|\psi_i\rangle, p_i\}$ according to the procedure given in Section 4.1.2, then one can go beyond the minimal interpretation – and claim that they are indeed really and objectively each in one of the states $|\psi_i\rangle$. However, subjectively we may not know in which one, i. e. we are ignorant. In principle, however, this can be known, for example if the possibility exists of determining precisely which individual preparation apparatus was active (see Fig. 4.1). One then says that the state ρ allows an *ignorance interpretation*. Of course, one could also prefer a minimal interpretation of quantum theory, in which the question posed above is never asked in the first place.

4.4 Density Operators of Qubits

The density operator ρ in $\mathcal{H}_2$ can be decomposed as in Sect. 3.1 in terms of the Pauli operator basis

$$\rho = \frac{1}{2}(\mathbb{1} + \mathbf{r}\boldsymbol{\sigma}) \; . \tag{4.61}$$

Here, $\boldsymbol{r}$ is the Bloch vector

$$\mathbf{r} = \mathrm{tr}[\rho\boldsymbol{\sigma}] = \langle\boldsymbol{\sigma}\rangle, \quad \mathbf{r} \in \mathbb{R} \; . \tag{4.62}$$

Using

$$\mathrm{tr}[\rho^2] = \frac{1}{2}(1 + |\mathbf{r}|^2) \tag{4.63}$$

as obtained from Eq. (3.23), it follows from

$$\frac{1}{2} \leq \mathrm{tr}[\rho^2] \leq 1 \tag{4.64}$$

that the Bloch vector obeys the inequality

$$|\mathbf{r}|^2 \leq 1 \; . \tag{4.65}$$

The degree of mixtureness is determined directly from the magnitude of the Bloch vector:

$$\Xi = \frac{1}{2}(1 - |\mathbf{r}|^2) \; . \tag{4.66}$$

For a true mixture, the Bloch vector $\mathbf{r}$ *lies within the Bloch sphere. The completely mixed state* $\frac{1}{2}\mathbb{1}$ *is represented by the centre of the sphere,* $\mathbf{r} = \mathbf{0}$.

Determination of states The relation $\rho \leftrightarrow \boldsymbol{r}$ as in Eqs. (4.61) and (4.62) is, in contrast to the relation $|\psi\rangle \leftrightarrow \mathbf{r}$, completely unambiguous, since phase factors are not represented by ρ. We can determine the state ρ by measurement of the expectation values $\langle\boldsymbol{\sigma}\rangle$ of the three different observables σ. In the case of the spin, this takes the form of a measurement of the mean values of the spin components in three independent directions. For a pure state, owing to the normalisation of $\mathbf{r}$, two directions suffice.

To conclude, we give the matrix elements as functions of the components of the Bloch vector:

$$\rho(\mathbf{r}) = \frac{1}{2}\begin{pmatrix} 1 + r_3 & r_1 - ir_2 \\ r_1 + ir_2 & 1 - r_3 \end{pmatrix} \tag{4.67}$$

with

$$\begin{aligned} r_3 &= \rho_{00} - \rho_{11} \\ r_2 &= i(\rho_{01} - \rho_{10}) \\ r_1 &= \rho_{01} + \rho_{10} \; . \end{aligned} \tag{4.68}$$

4.5 Complementary Topics and Further Reading

- Gleason's theorem: We initially described quantum states in terms of state vectors and thereafter by density operators. Are there other mathematical objects which permit unambiguous predictions of measurement probabilities? *Gleason's theorem* [Gle 57] states that

$$\langle A \rangle = \mathrm{tr}[A\rho] \tag{4.69}$$

  is the most general formula for the expectation value which is compatible with the probability structure of the quantum theory, if the dimension of the Hilbert space is greater than 2. The description of states by means of positive operators with trace 1 is the most general quantum-mechanical formulation. This is still true when measurements are described in terms of a POVM (cf. Sect. 13.4). Literature: [Per 93, pp. 190f], [BGL 95, p. 124f], [Aul 00, pp. 199f]. In [Bus 99], this theorem is proved also for the dimension 2.

- Complementary literature on density operators: [Fan 57].

- On the terminology "proper mixture": [d'Es 95] and [d'Es 99].

4.6 Problems for Chapter 4

Prob. 4.1 [for Sects. 4.1 and 4.4]: The state of a quantum system is given by the density matrix

$$\rho = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$$

with $a, b \in \mathbb{R}$, $a \geq 0$, $b \geq 0$ and $a + b = 1$.

a) What is the probability of obtaining the value $+1$ or -1 in a measurement of σ_x?

b) Calculate directly, without referring to a), the expectation value for a measurement of σ_x.

c) Compute the corresponding quantities for σ_y in place of σ_x.

d) Find the Bloch vector and explain the results for the expectation values.

Prob. 4.2 [for Sect. 4.3]: Give several different ensemble decompositions of the density matrix

$$\rho = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

for a spin system.

Prob. 4.3 [for Sect. 4.4]: Show that a density operator can be simply written as a function of the magnitude of the Bloch vector by making use of the eigenbasis:

$$\rho \leftrightarrow \begin{pmatrix} 1-|\mathbf{r}|^2 & 0 \\ 0 & 1+|\mathbf{r}|^2 \end{pmatrix}$$

(Hint: determine the eigenvalues of $\rho(\mathbf{r})$.)

Prob. 4.4 [for Sect. 4.4]: The observables A, B and C have the representations

$$A = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & -2i \\ 2i & 0 \end{pmatrix} \tag{4.70}$$

in the computational basis. Measurements on the state with the density operator ρ lead to the expectation values

$$\langle A \rangle = 2\,, \quad \langle B \rangle = \frac{1}{2}\,, \quad \langle C \rangle = 0\,. \tag{4.71}$$

Find the density operator ρ.

© 2007 WILEY-VCH Verlag GmbH & Co.

5 Shannon's Entropy and Classical Information

Entropy is a concept which was developed in the framework of thermodynamics. In classical statistical mechanics and in quantum statistics, it is used for the description of statistical mixtures. The entropy is in this case a measure of the disorder and of the missing information about a state. Beginning from this interpretation, the entropy developed into a key concept in classical information theory (*Shannon's entropy*) and quantum information theory (*von Neumann's entropy*).

Quantum information theory describes the transfer and processing of information using quantum systems as carriers of information. Here, von Neumann's entropy plays a multiple role:

(i) It permits statements about the classical information content coded on quantum-mechanical devices.

(ii) It quantifies which quantum-mechanical resources are at a minimum required in order to store a given amount of information.

(iii) Finally, it carries out a task which has no classical analogue: using the von Neumann entropy, the degree of entanglement of composite systems can be quantified.

In addition to these three tasks, there are other applications in connection with the theory of measurement, with perturbed quantum channels etc.

We start with the first point above, which follows immediately from our previous considerations of statical mixtures, and initially introduce as a preparation for the rest of this chapter the classical entropy of Shannon. Then, in the following chapter, we investigate information transport through quantum channels in order to give the quantum entropy an operational meaning. The significance of the concept of entropy for the description of entangled quantum systems will be sketched in the chapter after the next.

5.1 Definition and Properties

Formulation of the problem In this introductory section, we wish to explain the concept of Shannon's entropy using the example of a written text (e. g. a newspaper), which is intended to convey information. The text consists of a series of n letters from an alphabet.

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

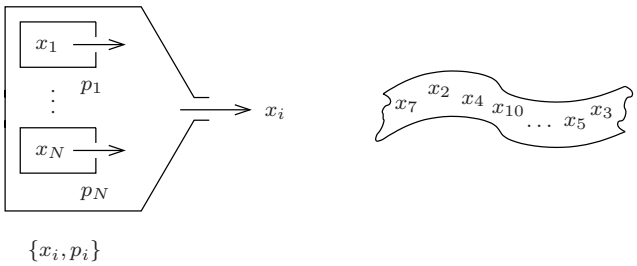


Figure 5.1: A signal source with the signal ensemble $\{x_i, p_i\}$ generates a character string.

Texts in a natural language are objects which are much too complex for our purposes. They consist for example of English words, i. e. not all possible combinations of letters are allowed. Furthermore, series of letters are not independent of each other, for example in an English text a 'q' is nearly always followed by 'u' but never by 'x'. We shall ignore these correlations between letters and consider only texts which are produced by a *signal source without memory*, and which meet one condition: a letter x_i is presumed to occur with the probability p_i, where $\sum_{i=1}^{N} p_i = 1$. N is the number of different letters, which depends on the language. In an English text, for example, the letter 'z' occurs less frequently than 'y', in contrast to German where the reverse is true. Such a *stochastic memoryless signal source* is characterised only by the *signal ensemble* $\{x_i, p_i\}$ with $i = 1, \ldots, N$. It represents a set of distinguishable alternatives x_i together with their probabilities p_i. The signal ensemble is also denoted as a *random variable* X.

As an operational implementation, one can imagine that N printing machines, each of which can print one letter x_i, are all working in parallel as one signal source. Using these printers, Alice, as sender, can now print out a text. This *message* consists of a *character string* or a *sequence* of n letters. The number of possible messages is then given by N^n. The printers are however set up in such a way that they fulfill a constraint. We consider a large number of such sequences as they are printed out by the source. The relative frequency with which the letter x_i is printed will be given by p_i. In the special case that $p_1 = 1$, there is e. g. only a single sequence: $x_1 x_1 x_1 \ldots x_1 \ldots$. But even when $p_1 \neq 1$, this sequence can occur[1]. All the machines together form a large printing device which represents the signal source in our imagined setup (see Fig. 5.1).

The particular text which is printed out by Alice is passed on to the recipient, Bob, letter for letter without errors using a carrier describable by classical physics (e. g. printed paper). Bob does not know which letters he will receive next, but he is supposed to know which printing device is used by Alice. *Bob thus has an important* advance information: *he knows the signal ensemble* $\{x_i, p_i\}$ *and thus in particular the probabilities* p_i. This is his *a priori* knowledge. We seek a measure of Bob's remaining *a priori* uncertainty given this knowledge. This measure can at the same time serve as a measure for the *information* which would remove

[1] With this model of printing machines, we have detached ourselves to a large extent from the "language", since e.g. $yy \ldots y$ will likewise be printed by a source whose p_i correspond to the frequencies of occurrence of letters in the English language, but this will to be sure happen only very rarely.

this *uncertainty*. As we shall see, the entropy represents such a measure of the information. This will already become clear in the following heuristic introduction of the entropy. A more precise quantification and operationalisation will then be discussed in the following Sect. 5.2.

Shannon's entropy [2] Alice prints out a sequence of n letters. There are N^n such sequences. If n is a large number, then it is probable that many of these sequences already reflect the relative frequencies p_i within themselves, that is the letter x_i turns up with the frequency $n_i = np_i$ somewhere within the long sequence. Sequences of the form $x_i x_i \dots x_i$ are not excluded but are improbable when the p_i are small. When n is large, Bob may therefore assume that he will receive one of the sequences which contain the letters x_i with the frequencies n_i. We will call such sequences *typical sequences* in Sect. 5.2.1. How many different sequences of this type are there? There are $n!$ ways to arrange n characters. Permutations of the same letters among each other does not produce a new text. For a character x_i, $n_i!$ permutations are possible. We therefore have

$$Z_n = \frac{n!}{n_1! n_2! \dots n_N!} \tag{5.1}$$

sequences with $\sum_{i=1}^{N} n_i = n$.

In order to arrive at probabilities p_i, we consider the limiting case of an infinitely long text, $(n \to \infty, n_i \to \infty)$. Then, $p_i = \frac{n_i}{n}$ and with Stirling's formula $\log(n!) = n \log n - n + \mathcal{O}(\log n)$ for the logarithm of the number Z_n, we obtain

$$\begin{aligned} \log Z_n &\to n \log n - n - \sum_{i=1}^{N} (n_i \log n_i - n_i) \\ &= -n \sum_{i=1}^{N} p_i \log p_i \ . \end{aligned} \tag{5.2}$$

(Here, we used $0 \log 0 = 0$.)

If we divide the logarithm of the number Z_n of possibilities by n, i.e. relate it to the individual characters as an average value, we obtain *Shannon's entropy* or the *classical entropy* $H(\tilde{p})$ of the *probability distribution* $\tilde{p} \leftrightarrow \{p_i, i = 1, \dots, N\}$, which is defined as follows:

$$H(\tilde{p}) := \lim_{n \to \infty} \frac{1}{n} \log Z_n = -\sum_{i=1}^{N} p_i \log p_i \geq 0 \ . \tag{5.3}$$

The logarithm is always computed to the base 2, since later, we wish to refer to bits. The notation $H(X)$ is often used instead of $H(\tilde{p})$. The number Z_n of possible texts with a number n of characters is then found in the limiting case to be

$$Z_n = 2^{nH(\tilde{p})} \ . \tag{5.4}$$

Since many (few) possibilities reflect a large (small) measure of *a priori* uncertainty for Bob, $H(\tilde{p})$ from Eq. (5.3) is a measure of the *mean* a priori *uncertainty* of a character which is

[2] [Sha 48], [Sha 49]

received by Bob. Then H is at the same time the average information which Bob receives per transmitted character. The associated overall information permits him to identify the sequence received from among the possible alternative sequences. H is dimensionless. The value of H gives the amount of information in units of *bits*. We will restate this more precisely in Sect. 5.3. *The entropy $H(\widetilde{p})$ thus characterises both the signal source and also the message.*

$H(\widetilde{p})$ is a function of the probability distribution p_i of the signal ensemble. Whether the signal consists of letters or of other symbols is irrelevant. The x_i can be any sort of alternatives which occur with the associated probabilities p_i. In contrast to signal transmission with quantum systems, we have assumed that the symbols (e. g. letters) are classical symbols. They can be uniquely distinguished from one another and are not changed in the process of reading out. If x_i is received by Bob, then he in fact reads x_i. Here, it plays no role with what sort of classical carrier (paper, tone frequencies, etc.) the information is transmitted. This independence on the physical properties of the classically-describable carrier shows that Shannon's entropy is a concept which is added on to classical physics rather than following from it.

Properties We first wish to prove some mathematical properties of the entropy. *The maximum value of $H(\tilde{p})$ is $\log N$, where N is the number of symbols in the signal ensemble. It is attained for the uniform distribution* $p_1 = p_2 = \ldots = p_n = N^{-1}$. To prove this conjecture, we write the constraint $\sum_i p_i = 1$ in the form $p_N = 1 - \sum_{i=1}^{N-1} p_i$ and consider the other $p_{i\neq N}$ as independent variables. The derivative of

$$H(\tilde{p}) = -\sum_{i=1}^{N-1} p_i \log p_i - p_N \log p_N \tag{5.5}$$

is then zero due to

$$\partial H/\partial p_l = -\log p_l + \log p_N \tag{5.6}$$

for $p_l = p_N = \frac{1}{N}$. This leads to

$$H^{\max}(\tilde{p}) = H(p_i = N^{-1}) = \log N \; . \tag{5.7}$$

There is no additional extremum, for example at the boundary. *We also note the minimum value of $H(\tilde{p})$*:

$$H^{\min}(\tilde{p}) = 0 \quad \Leftrightarrow \quad p_l = 1, \quad p_{i\neq l} = 0 \; . \tag{5.8}$$

It is obtained when the signal ensemble contains only a single symbol.

All together, we thus have

$$0 \leq H \leq \log N \tag{5.9}$$

and for *binary coding* $(N = 2)$:

$$0 \leq H \leq 1 \, . \tag{5.10}$$

$H(\widetilde{p})$ is in this case $(p_1 = p, \;\; p_2 = 1 - p)$ a function of p with $0 \leq p \leq 1$,

$$H(p) = -p \log p - (1 - p) \log(1 - p) \tag{5.11}$$

which is shown in Fig. 5.2. A classical system with two states has a maximum information capacity of $H(p = \frac{1}{2}) = 1$ bit.

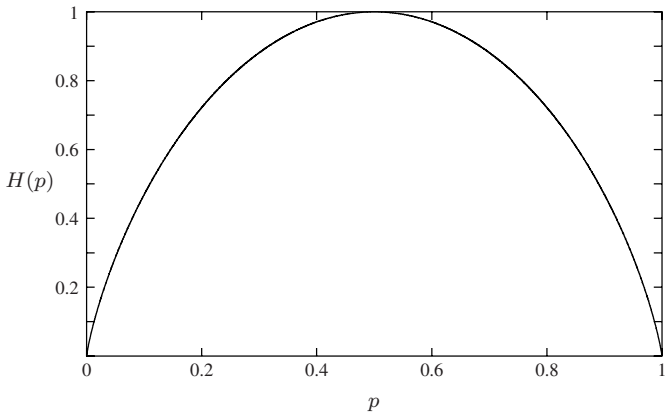


Figure 5.2: Shannon's entropy $H(p)$ for binary coding with the probabilities $p_1 = p$ and $p_2 = 1 - p$.

5.2 Shannon's Theorem

5.2.1 Typical Sequences

We now wish to show that the entropy $H(\widetilde{p})$ represents a reasonable measure of the information content per character in the case that a long text $(n \to \infty)$ is produced by a source with the signal ensemble $\{x_i, p_i\}$. How can one understand information *operationally* in a more precise way? Let us assume that Alice has a text, which we shall call the *output text*, and she wants to transmit it to Bob. Alice doesn't use the signal ensemble itself, but instead the simplest non-trivial alphabet. It consists of two characters (e. g. of the *binary digits* 0 and 1), which are assumed to occur with equal probabilities. The new text is then a *bit sequence* (or binary string). The transmission to Bob is again assumed to be free of errors. Each time that one of these symbols is received by Bob, we say that he has received the quantity of information 1 *bit*. The answer to the yes-no question, "Has the number 0 appeared?" contains the information 1 bit. We determine the information content of the output text by counting how many bits Alice must transmit in the most economical way, in order that Bob can determine which was the output text out of a set of texts of length n. Since we are considering very long texts, it suffices to determine the mean number of bits required per character of the output alphabet. We will show that it is equal to Shannon's entropy $H(\widetilde{p})$. The information in the output text is then $nH(\widetilde{p})$ bits. To analyse Alice's procedure, we start from the law of large numbers.

Limiting-case theorem We consider a stochastic, memoryless source, which produces not letters, but rather real numbers y_i , $i = 1 \ldots N$ with the probabilities p_i. The signal ensemble is thus $\{y_i, \psi_i\}$. A sequence of n of these numbers might be e. g.

$$y_4 y_1 y_{17} y_4 \ldots y_1 \,. \tag{5.12}$$

We enumerate the numbers in order. The real number at position l is termed w_l, with $l = 1, \ldots, n$. In (5.12), for example, $w_2 = y_1$. The sequence in (5.12) can then be written in the enumerated form

$$w_1 w_2 w_3 \ldots w_n \ . \tag{5.13}$$

We want to consider very long sequences. Then the *law of large numbers* applies. It states that the arithmetic average of the numbers in the sequence approaches the expectation value computed with the probabilities p_i:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{l=1}^{n} w_l = \sum_{i=1}^{N} p_i y_i =: \langle y \rangle \ . \tag{5.14}$$

The limiting case can be formulated somewhat more precisely: for given arbitrary small values $\epsilon > 0$ and $\delta > 0$, and assuming finite variance of the y_i, there is a large sequence length n such that the probability of generating a sequence with $\left| \frac{1}{n} \sum_l w_l - \langle y \rangle \right| < \delta$ is greater than $1 - \epsilon$.

Typical sequences After these preliminary considerations, we return to our letters or symbols x_i in the signal ensemble $\{x_i, p_i\}$. A particular sequence of length n is e. g.

$$x_4 x_1 x_{17} x_4 \ldots x_1 \ . \tag{5.15}$$

The total probability $P(x_4 \ldots x_1)$ for the occurrence of this sequence is the product

$$P(x_4 \ldots x_1) = p_4 \cdot p_1 \cdot p_{17} \cdot p_4 \ldots p_1 \ . \tag{5.16}$$

We take the negative logarithm and divide by n:

$$-\frac{1}{n} \log P(x_4, x_1, \ldots, x_1) = \frac{1}{n} \left\{ -\log p_4 - \log p_1 \ldots - \log p_1 \right\} \ . \tag{5.17}$$

This sum has the structure of an average value. The summands in brackets form a sequence of real numbers. It corresponds to the sequence (5.12) with $y_4 = -\log p_4$, etc. Referring to Eq. (5.13), we can write: $w_1 = -\log p_4, \ldots, w_n = -\log p_1$. The probability of occurrence of $-\log p_i$ on the right-hand side of Eq. (5.17) is equal to that for the occurrence of x_i in the sequence (5.15), and therefore equal to p_i. With Eq. (5.14), we can again write the arithmetic average as an expectation value. Equation (5.17) together with the definition (5.3) then leads to:

$$\lim_{n\to\infty} \left\{ -\frac{1}{n} \log P \right\} = -\sum_{i=1}^{N} p_i \log p_i =: H(\widetilde{p}) \ . \tag{5.18}$$

Again, we will make this result more precise by introducing infinitesimal values ϵ and δ. Higher powers of ϵ and δ will be neglected. The probability that a sequence occurs with $P(\ldots)$ which leads to a value of $-\frac{1}{n} \log P$ within the interval

$$H - \delta < -\frac{1}{n} \log P < H + \delta \ , \tag{5.19}$$

is greater than $1-\epsilon$. Therefore, a sequence will be generated with near certainty whose total probability P from Eq. (5.16) fulfills the condition (5.19). Or, formulated as above: For a given ϵ and δ there is a sequence length n such that (5.19) is fulfilled with the probability $1-\epsilon$. With increasingly smaller ϵ and δ values, n becomes greater and greater. We rearrange Eq. (5.19):

$$2^{-n(H+\delta)} \leq P \leq 2^{-n(H-\delta)} \; . \tag{5.20}$$

Equation (5.20) states that with near certainty (at least with a probability $1-\epsilon$), sequences occur whose total probabilities P computed as in Eq. (5.16) are all equal to 2^{-nH}. These sequences are called *typical sequences*. *For sufficiently large n, the set of sequences of length n can be decomposed into two disjoint subsets: the typical sequences which occur with identical probabilities, and the remainder consisting of atypical sequences.*

How large is the number $Z(n,\epsilon,\delta)$ of such typical sequences? The sum of the probabilities of all the typical sequences must lie between $1-\epsilon$ and 1:

$$1-\epsilon \leq Z(n,\epsilon,\delta)P \leq 1 \; . \tag{5.21}$$

We divide by P and take into account the fact that P itself lies in the interval defined in Eq. (5.20). *Then it follows for the number of typical sequences*:

$$(1-\epsilon)2^{n(H-\delta)} \leq Z(n,\epsilon,\delta) \leq 2^{n(H+\delta)} \; . \tag{5.22}$$

This specifies Eq. (5.4) more precisely.

5.2.2 Classical Data Compression

Coding of long blocks and data compression The inequalities (5.20) and (5.22) have an immediate practical application. With an increasing number n of symbols in the message, atypical sequences almost never occur. The number $Z(n,\epsilon,\delta)$ of typical sequences approaches 2^{nH}:

$$Z(n,\epsilon,\delta) \xrightarrow{n\gg 1} 2^{nH} \; . \tag{5.23}$$

Furthermore, all the typical sequences occur with the same probability $P = 2^{-nH}$ (uniform distribution). We enumerate the 2^{nH} different typical sequences by natural numbers in the binary system. We thus code entire sequences (*block coding*) and no longer consider individual signals. We then require numbers with nH digits (assuming $H \neq 0$).

On the other hand, the number of possible sequences overall is $N^n = 2^{n \log N}$. If we enumerate them with binary numbers, we require numbers with $n \log N$ digits. For the characterisation of a particular typical sequence of length n by block coding (when $H \neq H^{\max} = \log N$), a binary text of the shorter length nH (compare Eq. (5.9)) suffices. A still more compact coding with fewer bits is indeed not possible, since all the typical sequences are already equally probable. We would gain nothing by recoding. The binary text is now transmitted to the receiver. Since the block coding is known to the receiver, he can

uniquely reconstruct the particular typical sequence. *Through the limitation to typical sequences – and only such occur as messages when n is large – combined with binary enumeration of the typical sequences, we have accomplished a* data compression*. Further compression is not possible.* Referring to an analogous statement about quantum systems as data carriers, we note also: *Shannon's entropy gives the number of classical binary information carriers per symbol which is at a minimum necessary to transmit the information contained in a message.* Classical binary information carriers can be e. g. slips of paper on which either a 0 or a 1 is printed; or tones which are transmitted with only one of two possible frequencies, etc.

Shannon's theorem We formulate the result (5.22) once again in a different way and consider the error probability in the process. The right-hand inequality of (5.22) states that we can map each typical sequence uniquely in a string of $n(H + \delta)$ binary digits. The remaining less-probable atypical sequences are mapped by Alice "erroneously" all onto a single binary string (e. g. 000...00). Then it is possible with this procedure that two different original messages are coded by the same binary string and an *error* can occur on decoding. We write the error probability in the form $1 - F$. F is called the *fidelity of the coding-decoding scheme*. *Shannon's noiseless coding theorem*[3] summarises the previous considerations in the following form: *When* n$(H + \delta)$ *bits are available for the coding of messages of the (large) length n. then the messages can be coded with an error probability $1 - F < \varepsilon$ in the corresponding binary strings (for given ϵ and δ, there is an n for which this statement holds). If only $H - \delta$ bits are available, then the error probability is greater than $1 - \epsilon$.*

5.3 Classical Information

We now return to the problem which we formulated at the beginning of Sect. 5.1. Alice has at hand a text whose content she knows, with n characters, obtained from a memoryless source with the signal ensemble $\{x_i, p_i\}$. The corresponding entropy is $H(\widetilde{p})$. Alice has previously agreed with Bob, who knows the signal ensemble, how the typical sequences will be enumerated digitally with strings of length $nH(\widetilde{p})$. As we showed in Sect. 5.2.2, Bob must then ask Alice $nH(\widetilde{p})$ yes-no questions in order to ascertain what is the number of the output text and thus which of the texts is the output text itself. This immediately yields an *operational interpretation of the entropy: Shannon's entropy $H(\widetilde{p})$ is the average number of required yes-no questions per character (symbol) of the output text*. The smaller the value of Shannon's entropy $H(\widetilde{p})$, the fewer questions Bob needs to ask. It is therefore reasonable to call $H(\widetilde{p})$ the *information content per character of the output text*. If the signal ensemble $\{x_i, p_i\}$ with $i = 1, \ldots, N$ has e. g. only one single character ($N = 1$), then Bob already knows the text. He need ask no questions at all. This corresponds to $H(\widetilde{p}) = 0$ (cf. Eq. (5.7)). As shown by Eq. (5.8), Bob must ask the most questions per character – namely $\log N$ – when the probabilities p_i are uniformly distributed: $p_1 = p_2 = \ldots = p_N = N^{-1}$.

[3] [Sha 48], [Sha 49]

5.4 Classical Relative Entropy

Gibbs' inequality It will be useful for many proofs to introduce as a tool the *classical relative entropy* $H(\tilde{p}\|\tilde{q})$ of $\{p_i\}$ relative to $\{q_i\}$ for two probability distributions $\{p_i\}$ and $\{q_i\}$ of the same alphabet $\{x_i\}$, and to employ its properties:

$$H(\tilde{p}\|\tilde{q}) := \sum_{i=1}^{N} p_i \log \frac{p_i}{q_i} = -H(\tilde{p}) - \sum_{i=1}^{N} p_i \log q_i \ . \tag{5.24}$$

Making use of the fundamental inequality relating the logarithms for all positive x

$$\log x \ln 2 = \ln x \leq x - 1 \ , \tag{5.25}$$

we can derive an inequality for $H(\tilde{p}\|\tilde{q})$:

$$\begin{aligned} H(\tilde{p}\|\tilde{q}) &= -\sum_{i=1}^{N} p_i \log \frac{q_i}{p_i} \\ &\geq \frac{1}{\ln 2} \sum_{i=1}^{N} p_i (1 - \frac{q_i}{p_i}) \\ &= \frac{1}{\ln 2} \sum_{i=1}^{N} (p_i - q_i) = 0 \ . \end{aligned} \tag{5.26}$$

This relation is called *Gibbs' inequality. The relative entropy $H(\tilde{p}\|\tilde{q})$ is not negative. It is equal to zero if and only if $p_i = q_i$ holds for all i (identical distributions).*

Concavity The fact that Shannon's entropy is a concave function is a direct consequence of Gibbs' inequality. If $\tilde{p}_1$ and $\tilde{p}_2$ are two probability distributions with probabilities $\{p_{1i}\}$ and $\{p_{2i}\}$, then for the distribution $\tilde{p} = \lambda\tilde{p}_1 + (1-\lambda)\tilde{p}_2$ with the probabilities $\{p_i = \lambda p_{1i} + (1-\lambda)p_{2i}\}$ and $0 < \lambda < 1$, we have

$$H(\lambda\tilde{p}_1 + (1-\lambda)\tilde{p}_2) \geq \lambda H(\tilde{p}_1) + (1-\lambda)H(\tilde{p}_2) \ . \tag{5.27}$$

The equals sign holds when the distributions $\tilde{p}_1$ and $\tilde{p}_2$ are identical. *This means that if one averages over two probability distributions, then the entropy increases.* The proof follows from Gibbs' inequality (5.26) by writing out the terms using Eq. (5.24):

$$0 \leq \lambda H(\tilde{p}_1||\tilde{p}) + (1-\lambda)H(\tilde{p}_2||\tilde{p}) = H(\tilde{p}) - \lambda H(\tilde{p}_1) - (1-\lambda)H(\tilde{p}_2) \ . \tag{5.28}$$

5.5 Mutual Information as a Measure of the Correlation between Two Messages

In Chapter 7, we shall see that in measurements on a subsystem of a bipartite quantum-mechanical system, the results of the measurements occur with particular probabilities. Furthermore, one finds correlations between the results when pairs of measurements are carried

out separably on the two subsystems. Both statements depend on the composite state of the composite system and characterise this state. We wish to prepare for the quantum-mechanical concepts with which the corresponding details can be quantitatively described, by first introducing the analogous classical concepts.

We assume that two signal sources with signal ensembles $X \leftrightarrow \{x_i, p_i\}$ and $Y \leftrightarrow \{y_j, p_j\}$ are available, whose symbols or signals are not produced independently of one another. That is, there are correlations. We seek a measure of the correlation of messages from the two signal ensembles. For simplicity, we write for the ensembles $X \leftrightarrow \{x, p(x)\}$ and $Y \leftrightarrow \{y, p(y)\}$, as well as $\sum_i = \sum_x$, etc. We then have e. g. $H(\tilde{p}(x)) = H(X)$.

$p(y|x)$ is the conditional probability of the occurrence of y when x has already occurred. As we discussed in Sect. 1.3, we then have

$$p(y|x)p(x) = p(y, x) = p(x, y) \; . \tag{5.29}$$

$p(x, y)$ is here the probability that X and Y occur together (the joint probability). The *average uncertainty of the occurrence of a pair* for known signal ensembles $\{(x, y), p(x, y)\}$ is also called the *joint entropy* $H(X, Y)$. According to Eq. (5.3), it takes the form

$$H(X, Y) = -\sum_{x,y} p(x, y) \log p(x, y) = H(Y, X) \; . \tag{5.30}$$

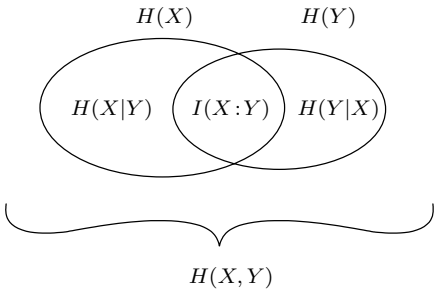


Figure 5.3: A set-theoretical visualisation of the different types of entropy.

5.5.1 Mutual Information

We introduce the quantity *mutual information* as

$$I(X:Y) := H(X) + H(Y) - H(X, Y) = I(Y:X) \tag{5.31}$$

with the entropies $H(X) := -\sum_x p(x) \log p(x)$ and $H(Y)$ for the individual ensembles (cf. Fig. 5.3). If we solve Eq. (5.31) for $H(X, Y)$,

$$H(X, Y) = H(X) + H(Y) - I(X:Y) \; , \tag{5.32}$$

we can see *that $I(X:Y)$ is a measure of how much less the uncertainty of the pairs is than the sum of the uncertainties of the two individual ensembles. This is a reasonable measure of the correlation of the two ensembles.*

We give an example. Alice has only blue and green socks. She always wears two socks of the same colour. Bob knows this and he also knows that Alice wears green socks with the probability $p(g) = \frac{1}{4}$ and blue socks with the probability $p(b) = \frac{3}{4}$. The ensemble $X \leftrightarrow L$ and $Y \leftrightarrow R$ refer to the left and to the right socks. Then we have $p(g,g) = \frac{1}{4}$, $p(b,b) = \frac{3}{4}$, $p(g,b) = 0$, $p(b,g) = 0$ and thus

$$H(L,R) = H(L) = H(R) = -0,25\log(0,25) - 0,75\log(0,75) = 0,81\ . \qquad (5.33)$$

Before Bob looks e. g. at the left sock, his uncertainty about the colour of the right sock is equal to $H(R)$. If, however, he looks at the left sock, then this uncertainty vanishes completely. Bob then has obtained the information

$$I(L\!:\!R) = H(L) + H(R) - H(L,R) = H(R)\ . \qquad (5.34)$$

Precisely this is reflected in the relation (5.31) together with Eq. (5.33). $I(L\!:\!R)$ is maximal, namely $I(L:R) = 1$, for $p(g) = p(b) = \frac{1}{2}$. The mutual information depends, like every information, on the previous knowledge. When we take it to be a measure of the correlation of the two signal ensembles, this point should be kept in mind.

5.5.2 Conditional Entropy

It is instructive to clarify the interpretation of $I(X:Y)$ once more in another way. We can interpret the entropy $H(X) = -\sum_x p(x)\log p(x)$ also as follows: $-\log p(x)$ is the uncertainty in the occurrence of the signal x. Weighting with the probability $p(x)$ of occurrence and summing over all x leads to $H(X)$ as the average uncertainty per signal. With the probability $p(x,y)$ that both x and also y occur, and the uncertainty $-\log p(x|y)$ for the occurrence of x when y has already occurred, we find analogously

$$H(X|Y) := -\sum_{x,y} p(x,y)\log p(x|y) \geq 0\ . \qquad (5.35)$$

This quantity is, in contrast to $H(X,Y)$ from Eq. (5.30), based on the conditional probability $p(x|y)$. $H(X|Y)$ is called the *conditional entropy*. *$H(X|Y)$ describes just how uncertain we still are (on average) concerning the value of x, when we already know the result y. We are*

thus dealing with a remaining uncertainty[4]. We rewrite $H(X,Y)$ with the aid of Eq. (5.29):

$$\begin{aligned} H(X,Y) &= -\sum_{x,y} p(x,y)\log p(x|y)p(y) \\ &= -\sum_{x,y} p(x,y)\log p(y) - \sum_{x,y} p(x,y)\log p(x|y) \\ &= -\sum_{y} p(y)\log p(y) + H(X|Y)\,. \end{aligned} \tag{5.38}$$

With this, we find

$$H(X|Y) = H(X,Y) - H(Y) \quad \text{and} \quad H(Y) \le H(X,Y)\,. \tag{5.39}$$

The uncertainty for pairs, $H(X,Y)$, is reduced to a remaining uncertainty $H(X|Y)$ when the information $H(Y)$ about the y signal is at hand. Analogous relations hold when the x signal has already been received. In our example of the socks, the remaining uncertainty is zero. A visualisation of the different entropy concepts is given in Fig. 5.3.

Equations (5.31) and (5.39) lead to

$$H(X|Y) = H(X) - I(X:Y)\,. \tag{5.40}$$

The uncertainty about x when y has occurred is reduced, compared to the a priori *uncertainty about the occurrence of x, when correlations between the ensembles X and Y are present. One can also interpret $I(X:Y)$ as the average information which one obtains about the value of X when the value of Y becomes known, and* vice versa *($I(X:Y) = I(Y:X)$). In this sense too, the mutual information $I(X:Y)$ is found to be a measure of the correlation of the ensembles.* We shall return to the mutual information in Chap. 9 in connection with entangled quantum systems.

Subadditivity Finally, we mention without proof that by making use of the convexity of the logarithm function, the inequality

$$H(X) \ge H(X|Y) \ge 0 \tag{5.41}$$

can be proven. Then it follows with Eq. (5.40) that the mutual information cannot be negative:

$$I(X:Y) \ge 0\,. \tag{5.42}$$

[4]The conditional entropy $H(X|Y)$ can also be introduced in a somewhat different manner. The average remaining entropy of the variables X, when the particular signal y has been received, is found from the conditional probability $p(x|y)$:

$$H(X|y) = -\sum_{x} p(x|y)\log p(x|y)\,. \tag{5.36}$$

Averaging over y leads with (5.29) to the conditional entropy $H(X|Y)$

$$H(X|Y) = -\sum_{y} p(y)\sum_{x} p(x|y)\log p(x|y) = -\sum_{x,y} p(x,y)\log p(x|y)\,. \tag{5.37}$$

Information about Y cannot reduce out knowledge of X and *vice versa*. Equality is found precisely when X and Y are independent. Equation (5.31) leads as an immediate consequence to the *subadditivity* of the entropy (compare Fig. 5.3):

$$H(X,Y) \leq H(X) + H(Y) \; . \tag{5.43}$$

5.6 Complementary Topics and Further Reading

See also Sect. 6.6.

- Review articles: [Weh 78], [Ste 98], [Ved 02].

5.7 Problems for Chapter 5

Prob. 5.1 [for 5.2]: List the possible messages of length $n = 4$ composed of the binary digits 0 and 1 (i. e. $N = 2$). Demonstrate how the number of typical sequences changes when p_0 approaches 1 and p_1 approaches 0. How does the probability of the typical sequences change? What happens on increasing the length n?

Prob. 5.2 [for 5.5.2]: Prove the subadditivity (5.43) using Gibbs' inequality (5.26).

Prob. 5.3 [for 5.5.2]: Prove the inequalities

$$H(X) \geq H(X|Y) \geq 0 \; . \tag{5.44}$$

Prob. 5.4 [for 5.5.2]: Show that the mutual information $I(X:Y)$ is zero if and only if X and Y are independent, i. e. when $p(x,y) = p(x)p(y)$.

© 2007 WILEY-VCH Verlag GmbH & Co.

6 The von Neumann Entropy and Quantum Information

6.1 The Quantum Channel and Quantum Entropy

The quantum channel We consider the following physical situation (cf. Tab. 6.1):

A classical signal source generates the letters of a character string, one after the other. As we have seen, this signal source can be described by an ensemble $\{x_i, p_i\}$ with $i = 1, \ldots, N$. The message is to be transmitted via a *quantum channel*. In this process, identical quantum objects (e. g. atoms of the same type with spin $\frac{1}{2}$, or photons) take on the role of carriers of the quantum-mechanical signal alphabet. For each character x_i, a preparation apparatus with the index i generates a quantum system in the *signal state* $|\psi_i\rangle$ and transmits it (compare Fig 4.1 with Fig. 5.1). The relationship between the character x_i and the state $|\psi_i\rangle$ is unambiguous. The entire setup is called the *quantum signal source*. By means of the preparation procedure, the classical information is thus coded in terms of pure quantum states. This first interface produces a statistical mixture of signal states with the density operator

$$\rho = \sum_{i=1}^{N} p_i |\psi_i\rangle\langle\psi_i| \tag{6.1}$$

in a Hilbert space $\mathcal{H}_d$ of dimension d. The density operator ρ (also called the state ρ) again refers to a preparation procedure. The associated ensemble is the *quantum signal ensemble* $\{|\psi_i\rangle, p_i\}$. It is important that we not require in general that the normalised state vectors $|\psi_i\rangle$ be orthogonal. Furthermore, the dimension d need not be the same as N; N can be e. g. larger. The transmitting quantum channel should be free of disturbances and isolated from outside influences. The quantum signal ensemble thus –for simplicity– remains unchanged within the channel.

At a second interface, the attempt is made to read out the information which was originally input into the signal source, via projective measurements. For this purpose, a detector observable D is measured. The orthonormal eigenstates $\{|d_m\rangle,\ m = 1, \ldots, d\}$ of the observables D

$$D|\mathrm{d}_m\rangle = d_m|d_m\rangle \tag{6.2}$$

form an ONB of $\mathcal{H}_d$. The associated eigenvalues d_m are assumed not to be degenerate. Then the correspondence between the measured values d_m and the states $|d_m\rangle$ after the measurement procedure is unambiguous. The probability of occurrence of a measured value d_m in a measurement on ρ is denoted by $p(d_m)$. Transmission of a signal via a quantum channel reflects the underlying scheme of quantum theory: at the beginning stands the preparation of a

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

Table 6.1: Classical information is coded as quantum information. A measurement again yields the classical information.

state and at the end, a measurement. The input consists of a sequence of classical signals with the Shannon entropy $H(\tilde{p})$, and the output is a sequence of measured values with the Shannon entropy $H(\tilde{p}(d))$.

To complete the description, we make note of the spectral decomposition of the density operator ρ:

$$\rho = \sum_{m=1}^{d} \lambda_m |m\rangle\langle m| \ , \quad \langle m|m'\rangle = \delta_{mm'} \tag{6.3}$$

with the eigenvectors $|m\rangle$ and the eigenvalues λ_m. The $\{|m\rangle,\ m = 1, \ldots, d\}$ form an ONB of $\mathcal{H}_d$, which is also called the *eigenbasis* of ρ.

The von Neumann entropy We first consider a special situation, in which the classical information which is input can again be read out without losses. The quantum system is to this end chosen so that the dimension d of $\mathcal{H}_d$ is the same as the number N of characters in the classical signal ensemble. At the first interface, by a suitable choice of the preparation procedure, corresponding to the character x_i, an eigenstate $|d_i\rangle$ of some detector observable D is generated (i. e. $|\psi_i\rangle = |d_i\rangle$)

$$\rho = \sum_{i}^{N} p_i |\mathrm{d}_i\rangle\langle \mathrm{d}_i| = \sum_{i}^{N} \lambda_i |i\rangle\langle i| \ . \tag{6.4}$$

In this case, we thus have $p_i = \lambda_i$ and $|d_i\rangle = |i\rangle$. The quantum signal source becomes a quasi-classical source due to the distinguishability of the signal states. Subsequently, at the second interface, the observable D is measured. The occurrence of the value d_i gives a unique indication of the original input of the signal character x_i, owing to their distinguishability. All of the probability distributions involved are the same: $p(d_i) = p_i = \lambda_i$. Correspondingly, we obtain for Shannon's entropy of the signal ensembles and of the ensemble of the measured values the value $H(\tilde{p}) = H(\tilde{p}(d))$.

The unique relation between the ensembles $\{x_i, p_i\}$, $\{|\psi_i\rangle, p_i\}$, and $\{d_i, p(d_i)\}$ in this particular quasi-classical situation and the corresponding agreement of the three probability distributions suggest that we associate to the statistical mixture with the density operator ρ of Eq. (6.4) a *quantum entropy* $S(\tilde{\lambda})$ which has the same value as Shannon's entropy ($S(\tilde{\lambda}) = H(\tilde{p})$):

$$S(\tilde{\lambda}) = -\sum_{i=1}^{d} \lambda_i \log \lambda_i \geq 0 \ . \tag{6.5}$$

Using the spectral decomposition of ρ in Eq. (6.3), $S(\tilde{\lambda})$ can be written as a function of the density operator ρ:

$$S(\rho) := S(\tilde{\lambda}) = -\mathrm{tr}[\rho \log \rho] \geq 0 \ . \tag{6.6}$$

This quantum entropy $S(\rho)$ is also called the *von Neumann entropy*[1] *of the mixture with the density operator* ρ. The unit of this entropy is a quantum bit or a *qubit*. We shall give the reason for this notation in the next chapter.

[1] [vNe 68]

Since $S(\rho)$ is unambiguously determined when ρ is fixed, we can generalise from the special information-transmission procedure described above and associate formally a von Neumann entropy $S(\rho)$ as in Eq. (6.6) to every density operator ρ and thus to every quantum state, even in physical situations in which there is no signal transmission or processing. $S(\rho)$ characterises a density operator ρ independently of how the corresponding state was prepared physically. ρ can also be a reduced density operator, which describes the state of a subsystem of a multipartite system. A state ρ with a spectral decomposition (6.3) cannot be distinguished from a statistical mixture of the states of the eigenbasis with the ensemble $\{|m\rangle, \lambda_m\}$. The state ρ can thus be completely simulated in this way. With this statistical mixture, if it is generated as a signal source, the classical information $H(\tilde{\lambda}) = S(\rho)$ will on the average be transmitted per signal state. If one wishes to gain a descriptive understanding of the von Neumann entropy by association with the familiar Shannon entropy, then it is necessary to follow this circuitous route via the signal ensemble of the equivalent statistical mixture. In the next section, we shall see that an alternative visualisation is also possible.

Measurement of the von Neumann entropy For this purpose, it suffices to determine ρ and its eigenvalues. This is especially simple when the eigenstates $\{|m\rangle\}$ of ρ are already known. Measurements in the eigenbasis of ρ are particularly favourable. They lead to the probabilities

$$p(d_m) = \lambda_m \,, \tag{6.7}$$

with which the von Neumann entropy of ρ can be found:

$$S(\rho) = -\sum_{m=1}^{d} p(d_m) \log p(d_m) = H(\tilde{p}(d)) \,. \tag{6.8}$$

6.2 Qubits as the Unit of Quantum Information

The term *quantum information* is intended to refer to information which is represented by the state of a quantum system and which can be transmitted using the quantum system as carrier. In this case, the transmitter is a preparation apparatus and the receiver is a measurement device. These are the interfaces for the transitions classical $\rightarrow$ quantum-mechanical or quantum-mechanical $\rightarrow$ classical information. Only classical information is a type of information which we can read out directly.

It is our goal, in analogy to our procedure in Sect. 5.2, to transmit a sequence of quantum states $|\psi_1\rangle, \ldots, |\psi_n\rangle$ from a given signal ensemble using as few quantum systems as possible. It can be shown [Sch 95] that such a sequence can be compressed *without reference to classical information* in a unitary manner, so that it can be regenerated with asymptotically perfect fidelity by the receiver using a further unitary transformation. For this purpose, qubit systems are employed as quantum-mechanical carrier systems. *For optimal compression, the average number of qubit systems required per signal state is given by the von Neumann entropy of the signal ensemble described by the density operator,* $\{|\psi_i\rangle, p_i\}$. The von Neumann entropy thus obtains an operational interpretation, which no longer relies on the classical Shannon entropy. This theorem is the quantum-mechanical analogue of Shannon's theorem from Sect. 5.2.2. It

was derived by Schumacher in its most precise form using *quantum data compression* and is also called *Schumacher's quantum noiseless coding theorem*. We can not prove it at this point, since in the proof, the sequence of quantum states is taken to be a product state of a large composite system. Such quantum systems will be introduced in the next chapter. Concerning the proof, which is not very simple,[2] we refer to the literature (see Sect. 6.6). The unit of the quantum entropy $S(\rho)$ is termed the *qubit*. A quantum system with two levels (e. g. $|0\rangle$ and $|1\rangle$) which permits the coding of just one qubit of information is also itself called a qubit (qubit system would be preferable).

We now wish to compare the classical transmission of quantum information with the quantum-mechanical transmission. Let the signal ensemble consist of the states

$$|\psi_0\rangle = |0\rangle\,, |\psi_1\rangle = |0_x\rangle \tag{6.9}$$

which occur with the probabilities

$$p_1 = p_2 = \frac{1}{2}\,. \tag{6.10}$$

In a classical transmission of the information, owing to $H(p_1, p_2) = 1$ bit, no data compression is possible. For a sequence of n quantum states, after recoding we require n classical carriers of 1 bit.

It is more efficient not to recode to classical carriers. The von Neumann entropy is found from the density operator:

$$\rho = p_0|\psi_0\rangle\langle\psi_0| + p_1|\psi_1\rangle\langle\psi_1| = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \\ \frac{1}{4} & \frac{1}{4} \end{pmatrix}\,. \tag{6.11}$$

The matrix is defined in terms of the computational basis. The eigenvalues can be determined ($\lambda_0 = \cos^2\frac{\pi}{8} = 0,853$, $\lambda_1 = \sin^2\frac{\pi}{8} = 0,146$) and lead with Eq. (6.5) to $S(\rho) = 0,601$ qubits. For the transmission of a sequence of states of the signal ensemble, we therefore require only $0,601$ qubit systems per state. *This shows that a quantum compression with quantum coding in terms of qubit systems is a useful tool for the transmission of quantum information.* The qubit is an appropriate unit of quantum information. We should also emphasize that the compression procedure of Schumacher is universal. In order to carry it out, one need not know the state which is to be transmitted.

[2] The proof is similar to that of Shannon in Sect. 5.2, which is based on the idea of typical sequences. If we restrict ourselves to qubit systems, then the signal states $|\psi_i\rangle$ are states in $\mathcal{H}_2$ that need not be orthogonal. We combine the n systems of a long sequence into a composite system with a state vector in the 2^n-dimensional product space. Then the following can be shown (see Sect. 6.6): When the signal ensemble $\{|\psi_i\rangle, p_i\}$ has a von Neumann entropy $S(\rho) > 1$, then the probability (which increases with n) is large that the state vector lies within a *typical subspace* of the product space, depending on the ensemble. Its dimension is $2^{nS(\rho)}$. Therefore, product states with $nS(\rho)$ qubit states are sufficient for the representation of such state vectors. Correspondingly, only $nS(\rho)$ qubit systems as carriers are required for the transmission. From their composite state, the original sequence can be reconstructed via a unitary transformation. The error in this process decreases with increasing n.

6.3 Properties

Referring to Eq. (6.8), in analogy to Shannon's entropy, we can show directly that:

(i) A pure state $\rho = |\psi\rangle\langle\psi|$ has the minimal value of the entropy, $S(\rho) = 0$.

(ii) For a density operator with d non-vanishing eigenvalues, one finds

$$0 \leq S(\rho) \leq \log d \,. \tag{6.12}$$

The equals sign holds when all non-vanishing eigenvalues are the same. The completely mixed state $\rho = \frac{1}{d}\mathbb{1}$ in a Hilbert space of dimension d has the maximum von Neumann entropy, $S(\rho) = \log d$.

(iii) As a result of the concavity of Shannon's entropy, we find for the von Neumann entropy for $p_j > 0$ with $\sum_j p_j = 1$ the concavity relation

$$S(p_1\rho_1 + \ldots + p_r\rho_r) \;\geq\; p_1 S(\rho_1) + \ldots + p_r S(\rho_r) \,. \tag{6.13}$$

For the proof, one makes use of the spectral decomposition of ρ_j. $\sum_j p_j\rho_j$ is the state of a quantum system which is found in the unknown state ρ_j with the probability p_j. The result (6.13) is plausible. Several ensembles have been combined, or the associated ensemble decompositions have been mixed. Our lack of knowledge about this mixture is greater than our average lack of knowledge about the states ρ_j. The information about which mixture a state of the ensemble decomposition has come from has been lost. The entropy is greater because we know less about the preparation. It follows in particular from (i) that

$$S(\rho) > 0 \,, \tag{6.14}$$

if ρ is not a pure state.

Unitary dynamics It holds for a unitary transformation of the density operator that

$$S(U\rho U^\dagger) = S(\rho) \,, \tag{6.15}$$

since S depends only on the eigenvalues of ρ. The entropy is thus – independently of which representation one chooses for the unitary dynamic evolution – always independent of time:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = 0 \,. \tag{6.16}$$

Our information about a state does not change during the unitary dynamic evolution. The corresponding conclusion does not hold for the measurement dynamics.

Quantum-mechanical relative entropy and Klein's inequality This will serve us primarily as a mathematical computational aid. We consider two density operators ρ and σ, and introduce the *quantum-mechanical relative entropy* $S(\rho\|\sigma)$ of ρ relative to σ:

$$S(\rho \parallel \sigma) := \mathrm{tr}[\rho \log \rho] - \mathrm{tr}[\rho \log \sigma] \,. \tag{6.17}$$

As in the analogous classical case, we wish to derive an estimate. The orthogonal decompositions of ρ and σ are assumed to be given by

$$\rho = \sum_{m=1}^{d} \lambda_m |\phi_m\rangle\langle\phi_m| \,, \quad \sigma = \sum_{m=1}^{d} \kappa_m |\xi_m\rangle\langle\xi_m| \,. \tag{6.18}$$

It then follows that:

$$S(\rho\|\sigma) = \sum_m \lambda_m \log \lambda_m - \sum_m \langle\phi_m|\rho \log \sigma|\phi_m\rangle \,. \tag{6.19}$$

With $\langle\phi_m|\rho = \lambda_m\langle\phi_m|$, we can rewrite the second term to give

$$\langle\phi_m| \log \sigma|\phi_m\rangle = \langle\phi_m|(\sum_{m'} \log \kappa_{m'} |\xi_{m'}\rangle\langle\xi_{m'}|)|\phi_m\rangle = \sum_{m'} P_{mm'} \log \kappa_{m'} \,. \tag{6.20}$$

Here, we have introduced

$$P_{mm'} := \langle\phi_m|\xi_{m'}\rangle\langle\xi_{m'}|\phi_m\rangle \tag{6.21}$$

with the properties $P_{mm'} \geq 0$, $\sum_m P_{mm'} = 1$ and $\sum_{m'} P_{mm'} = 1$. We again rewrite Eq. (6.19):

$$S(\rho\|\sigma) = \sum_m \lambda_m (\log \lambda_m - \sum_{m'} P_{mm'} \log \kappa_{m'}) \,. \tag{6.22}$$

The logarithm is a concave function, therefore $\sum_{m'} P_{mm'} \log \kappa_{m'} \leq \log \mu_m$ with $\mu_m := \sum_{m'} P_{mm'} \kappa_{m'}$. Using Eq. (6.22), this inequality can be transferred:

$$S(\rho\|\sigma) \geq \sum_m \lambda_m \log \frac{\lambda_m}{\mu_m} \,. \tag{6.23}$$

The right-hand side is formally the same as a classical relative entropy. This leads to *Klein's inequality*:

$$S(\rho\|\sigma) \geq 0 \,. \tag{6.24}$$

The quantum-mechanical relative entropy is non-negative. It is zero if and only if $\rho = \sigma$ (identical states). Like its classical counterpart, we shall use this theorem primarily as a computational aid.

6.4 The Interfaces of Preparation and Measurement

In Section 6.1, our preparation procedure was ideally matched, as was our readout. Departures from this ideal matching lead to loss of information. What is the cause of this and how can we express it quantitatively?

In the transport and processing of quantum-mechanically coded information, three typical features of quantum theory are particularly important, which are lacking in classical physics: first of all the situation that non-orthogonal pure states cannot be perfectly distinguished by a measurement. Even orthogonal states can only then be distinguished when they are eigenstates of the observable operator. Secondly, a quantum-mechanical measurement in general modifies the state. A third point is the ambiguity of the ensemble decomposition of a density operator. This has the converse result that there are many classical ensembles with different Shannon entropies which lead to the same density operator after coding at the first interface, and then can no longer be distinguished by means of a measurement. The von Neumann entropy of the state ρ is determined by its orthogonal decomposition, which codes only precisely one of the classical ensembles. It is helpful to recall (cf. Sect. 5.1) that the entropy has two aspects. It characterises the remaining *a priori* uncertainty, and it is a measure of the quantity of information which will remove this uncertainty. We want to discuss the consequences of this in more detail and begin with the measurement process, that is the second interface of Tab. 6.1.

6.4.1 The Entropy of Projective Measurements

The states in the quantum channel are described by the density operator ρ with the von Neumann entropy $S(\rho)$. A *non-selective* measurement of the decoding observables D leads to a probability distribution $\{p(d_m)\}$ of the measured values d_m. One can conceive of the measurement as a classical stochastic source with the signal ensemble $\{d_m, p(d_m)\}$. The signal ensemble has the Shannon entropy $H(\tilde{p}(d_m))$. Here, $\{p(d_m)\}$ is at the same time the probability distribution of the states $\{|d_m\rangle\}$, into which the system is transformed by a non-selective measurement. As the result, a density operator ρ' with the orthogonal decomposition

$$\rho' = \sum_{m=1}^{d} p(d_m)|d_m\rangle\langle d_m| = \sum_{m=1}^{d} P_m \rho P_m \tag{6.25}$$

is obtained. For the projection operators, we have $P_l^2 = P_l$, $P_l P_m = \delta_{l,m} P_l$, and $\sum_l P_l = \mathbb{1}$. *Shannon's entropy $H(\tilde{q})$ of the measured values and the von Neumann entropy $S(\rho')$ of the mixture of the quantum states after the measurement are the same*:

$$S(\rho') = H(\tilde{p}(d))\,. \tag{6.26}$$

Klein's inequality permits us to compare the quantum entropies $S(\rho)$ and $S(\rho')$ before and after a non-selective measurement. We start with

$$0 \leq S(\rho\|\rho') = -S(\rho) - \mathrm{tr}[\rho \log \rho'] \tag{6.27}$$

and consider the second term more carefully:

$$\mathrm{tr}[\rho \log \rho'] = \mathrm{tr}[(\sum_l P_l)\rho \log \rho'] = \mathrm{tr}[\sum_l P_l \rho \log(\rho') P_l]\,. \tag{6.28}$$

We have made use of the properties of the projection operators and permuted the terms within the trace. Equation (6.25) shows that $P_l\rho' = P_l\rho P_l = \rho' P_l$ holds. Therefore, P_l also commutes with the operator function $\log\rho'$ and we find

$$\begin{aligned} \mathrm{tr}[\rho\log\rho'] &= \mathrm{tr}[\sum_m P_m\rho P_m\log\rho'] \\ &= \mathrm{tr}[\rho'\log\rho'] = -S(\rho')\,. \end{aligned} \tag{6.29}$$

Thus after inserting into Eq. (6.27) and making use of Eq. (6.26), we obtain the overall result:

$$S(\rho') \geq S(\rho)\,. \tag{6.30}$$

We compare two quantum entropies. *In a non-selective projective measurement, the von Neumann entropy of the state ρ' after the measurement is the same as the von Neumann entropy of the state ρ before the measurement if and only if the measurement takes place in the eigenbasis of ρ; otherwise, it is greater.* A non-selective measurement therefore transforms the system in general into a new signal ensemble with a greater entropy and in this manner it destroys information. The *a priori* uncertainty has increased as a result of the non-selective measurement.

This can be demonstrated with a very simple example. The pure state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \tag{6.31}$$

has a vanishing entropy. Non-selective measurement in the eigenbasis leads to the totally-mixed state

$$\rho = \frac{1}{2}\mathbb{1} \tag{6.32}$$

with maximum entropy, $S(\rho) = 1$.

6.4.2 The Entropy of Preparation

In quantum coding, the classical signal source with the ensemble $\{x_i, p_i\}$ is coded into the quantum ensemble $\{|\psi_i\rangle, p_i\}$, described by the density operator ρ of Eq. (6.1) (cf. Tab. 6.1). Due to the ambiguity of the ensemble decomposition of ρ, various classical ensembles with many different values of entropy H lead to quantum ensembles with the same density operator ρ and thus to the same von Neumann entropy $S(\rho)$. The maximum value of the quantum entropy is given via Eq. (6.12) by the dimension of the Hilbert space of the quantum system into whose states the coding takes place. The maximum value of the *preparation entropy* $H(\tilde{p})$ is given by the number N of characters in the classical alphabet x_i. It is the same as the number of state vectors $|\psi_i\rangle$, which can be larger than the number of basis vectors of $\mathcal{H}_d$, since the $|\psi_i\rangle$ need not be orthogonal. We thus expect a relation of the form

$$H(\tilde{p}) \geq S(\rho)\,. \tag{6.33}$$

This can indeed be verified in a long proof (see [Weh 78, p. 238] or [CD 94, p. 527]). In Eq. (6.33), equality is found if and only if the states $|\psi_i\rangle$ are mutually orthogonal. *If the signal*

states are not orthogonal, they cannot be distinguished. There is no decoding observable with which the full information content of the coded classical message could be read out again. ρ transmits less information than was contained in the original classical signal. Even with an optimally-matched final measurement, the information can no longer be completely recalled.

6.5 Quantum Information

We summarise:
In classical information theory, it is not important just how the carrier of the information is physically implemented. Printed characters can for example be converted in an error-free manner into the phonemes of spoken characters, and *vice versa*. As we have seen, quantum information can in general not be converted into classical information and back without losses. The cause of this lies among other things in the non-classical structure of the measurement process. *Quantum information is therefore generally a very different sort of information from classical information, just as a quantum state is a different sort of state from that of a classical system.*

Quantum information is stored in quantum states. Its carriers are quantum systems. Its transmission is accomplished by the propagation of the carriers between the preparation apparatus and the measurement apparatus. The processing of quantum information consists of the manipulation of quantum states. Unitary transformations and measurements are examples of this. The two types of information are associated with two different units of information: bits and qubits. Quantum information theory applies uniformly to all the different qubit systems (spins, photon polarisation states, etc.).

As we shall show in detail in the following chapters, the storage and processing of quantum information differs in essential ways from those of classical information: (i) the states of a qubit are not limited to 0 and 1. They are described by the entire Bloch sphere. (ii) The state of a quantum system which is a composite of several qubit systems can be entangled. (iii) Classically, there are only jumps between 0 and 1. Unitary transformations and other operations are however much more general and comprehensive. (iv) However, in a measurement, the quantum-mechanical final state cannot be simply read out like the classical state.

6.6 Complementary Topics and Further Reading

See also Sect. 5.6.

- Review articles: [Weh 78], [CD 94], [CF 96], [Ste 98], [Joz 98], [Ved 02].
- On the concept of “quantum information”: [Wer 01], [Wer 06].
- On Schumacher’s theorem and quantum data compression: [JS 94], [Ben 95], [Sch 95], [Joz 98], [Ved 02].

6.7 Problems for Chapter 6

Prob. 6.1 [for 6.1]: Determine the entropy of a state ρ in $\mathcal{H}_2$ as a function of the Bloch vector, with reference to a result from Chap. 3.

Prob. 6.2 [for 6.2]: Find the Bloch vector $\mathbf{r}$ corresponding to the density operator ρ of Eq. (6.11). The eigenvectors of ρ and $\boldsymbol{r\sigma}$ are the same (why?). Read off the representation of the eigenvectors in the computational basis with reference to Sect. 3.2 for $\mathbf{r\sigma}$, and find the eigenvalues λ_0 and λ_1. Compute $S(\rho)$.

© 2007 WILEY-VCH Verlag GmbH & Co.

7 Composite Systems

We turn now to composite systems, and begin by providing the necessary mathematical tools. We then generalise the postulates and discuss the special case of measurements on subsystems in detail. The consequences of entanglement will become clear in this discussion. We will demonstrate a conjuring trick which cannot be explained by classical physics. The unitary dynamics can once again be formulated with the aid of Liouville operators. The action of simple quantum gates on multiple qubit systems will be introduced.

7.1 Subsystems

We are accustomed from classical physics to the fact that *composite systems* (or compound systems) can be decomposed into their *subsystems* and that conversely, individual systems can be combined to give overall composite systems. The classical total system is completely describable in terms of the states of its subsystems and their mutual dynamic interactions. The solar system with the sun, the planets and the gravitational field is an example. In quantum physics, however, it is found that composite systems can have in addition completely different and surprisingly unified properties. These come to light when the composite quantum systems are in *entangled states* In such cases, it is indeed true in a certain sense that "the whole is more than the sum of its parts". We will present the details in a similar fashion as in Sect. 1.2 and begin with a discussion of preparation and measurements.

But first: what are composite systems? There are particular quantum systems which exhibit an internal structure. *One can distinguish in them two or more subsystems which can be accessed separately.* With this we mean that subsystems can be experimentally identified on which individually (and in this sense locally) interventions can be carried out. The corresponding operations are referred to as *local operations* These can be for example preparations or measurements.

We list some *bipartite systems* consisting of two subsystems. One can prepare quantum systems for which, in a measurement at two different locations, a photon can be registered at each location. There are analogous systems involving electrons. There are systems in which at one location a photon and at another an atom are detected. Subsystems are in general termed *local*, but they need not in fact be spatially separated. A composite system can be composed of e. g. an orbit (an external degree of freedom) and the polarisation (an internal degree of

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

freedom) of a single quantum object. Of course two separate systems, which are completely independent of one another, can also be considered formally as a total system.

It is essential not to assume e. g. for a 2-photon system that the photons involved are themselves distinguishable (which they are *not*, as is well-known). The *locations* at which for example the photon polarisation is measured *are* distinguishable. We know that in measurements on this system, always exactly two photons are prepared together and therefore the overall system is a bipartite system. The corresponding subsystems S^A and S^B are in this case associated with the locations of the detectors, A and B (a photon at the location A or a photon at the location B). In general, apparatus which carry out operations are classical objects and thus have an individual identity. In contrast, owing to the indistinguishability of the photons, the question of *which* photon was detected in a particular measurement, e.g. by the detector at the location A, makes no sense. We will return to this point in Sect. 7.9.

Alice and Bob In order to make it especially clear that measurements or manipulations are carried out on different subsystems S^A and S^B of the composite system S^{AB}, one often introduces the experimentalists *Alice* and *Bob*, who carry out local operations on the subsystem S^A or S^B (often, but not necessarily, at different locations). By referring to Alice and Bob, we emphasize once more that many quantum-mechanical statements are to be understood *operationally* (i. e. as instructions for carrying out an action); e.g. of the type: "If Alice does *this* to subsystem S^A, then Bob will measure *that* on subsystem S^B".

Existence We will once again assume, in agreement with the standard interpretation from Sect. 1.2, that such subsystems are not just abstract auxiliary constructions like the quantum systems in the minimal interpretation, but rather that they exist in reality. With this, we do not mean to imply that a state can be ascribed to an individual subsystem which is independent of the state of the other subsystem. In entangled systems, precisely this independence does not exist. This is the cause of many startling quantum-physical effects. It is furthermore not meant by our assumption of existence in reality that similar elementary particles of the same type, such as two photons, have individual identities and are therefore distinguishable. The assumption that the photons exist cannot lead us to such conclusions. *The possibility of separate manipulations, and not the individuality of quantum objects, defines the subsystem (compare Sect. 7.9).*

7.2 The Product Hilbert Space

We first wish to supply the mathematical formalism which we need to formulate the physics of composite systems. We require for this purpose the *product Hilbert space*.

7.2.1 Vectors

The *tensor product* $\mathcal{H}^{AB}$ of two Hilbert spaces $\mathcal{H}^A$ and $\mathcal{H}^B$, whose dimensions need not be the same,

$$\mathcal{H}^{AB} = \mathcal{H}^A \otimes \mathcal{H}^B \tag{7.1}$$

is itself a Hilbert space. We call $\mathcal{H}^A$ and $\mathcal{H}^B$ the *factor spaces*. For each pair of vectors $|\varphi^A\rangle \in \mathcal{H}^A$ and $|\chi^B\rangle \in \mathcal{H}^B$, there is a *product vector* in $\mathcal{H}^{AB}$, which can be written in different ways

$$|\varphi^A\rangle \otimes |\chi^B\rangle =: |\varphi^A\rangle|\chi^B\rangle =: |\varphi^A, \chi^B\rangle =: |\varphi, \chi\rangle \; . \tag{7.2}$$

It is linear in each argument with respect to multiplication by complex numbers.
With $\lambda, \mu \in \mathbb{C}$

$$|\varphi^A\rangle \otimes (\lambda|\chi_1^B\rangle + \mu|\chi_2^B\rangle) = \lambda|\varphi^A\rangle \otimes |\chi_1^B\rangle + \mu|\varphi^A\rangle \otimes |\chi_2^B\rangle \; , \tag{7.3}$$

and

$$(\lambda|\varphi_1^A\rangle + \mu|\varphi_2^A\rangle) \otimes |\chi^B\rangle = \lambda|\varphi_1^A\rangle \otimes |\chi^B\rangle + \mu|\varphi_2^A\rangle \otimes |\chi^B\rangle \; . \tag{7.4}$$

Entangled vectors If $\{|n^A\rangle\}$ is a basis of $\mathcal{H}^A$ and $\{|i^B\rangle\}$ is a basis of $\mathcal{H}^B$, then

$$\{|n^A\rangle \otimes |i^B\rangle\} \tag{7.5}$$

is a basis of $\mathcal{H}^{AB}$. For the dimension of $\mathcal{H}^{AB}$, we have $\dim\mathcal{H}^{AB} = (\dim\mathcal{H}^A)\cdot(\dim\mathcal{H}^B)$. Every vector $|\psi^{AB}\rangle$ in $\mathcal{H}^{AB}$ can be expanded in terms of the basis

$$|\psi^{AB}\rangle = \sum_{n,i} \alpha_{ni} |n^A, i^B\rangle \; . \tag{7.6}$$

All the definitions and statements can be directly applied to the product of a finite number of Hilbert spaces $\mathcal{H}^{AB\ldots M} = \mathcal{H}^A \otimes \mathcal{H}^B \otimes \cdots \otimes \mathcal{H}^M$. We introduce also the abbreviations:

$$\mathcal{H}^{\otimes n} := \mathcal{H} \otimes \mathcal{H} \otimes \cdots \mathcal{H}, \quad |\phi\rangle^{\otimes n} := |\phi\rangle|\phi\rangle \cdots |\phi\rangle \; . \tag{7.7}$$

Vectors in $\mathcal{H}^{AB}$ which are not product vectors are called *entangled*. They can be written only as a *superposition* of product vectors. We will represent entangled pure states with such vectors; they will play an important role in the following sections. The superposition is an important reason for this. It can usually not be read directly off the decomposition in terms of the basis (7.6) whether or not a vector $|\psi^{AB}\rangle$ is entangled. Later, we will develop a criterion (Sect. 8.3.1) and also extend the concept of entanglement to density operators (Sect. 8.1.1).

The scalar product The bra vector of the product vector $|\varphi^A\rangle \otimes |\chi^B\rangle$ has the form

$$(|\varphi^A\rangle \otimes |\chi^B\rangle)^\dagger = \langle\varphi^A| \otimes \langle\chi^B| =: \langle\varphi^A|\langle\chi^B| =: \langle\varphi^A, \chi^B| =: \langle\varphi, \chi| \; . \tag{7.8}$$

It follows from this for the dual corresponding vector of $|\psi^{AB}\rangle$ as in Eq. (7.6)

$$\langle\psi^{AB}| = \sum_{n,i} \alpha_{ni}^* \langle n^A, i^B| \; . \tag{7.9}$$

The scalar product is formed in a "space by space" manner:

$$\langle\varphi^A, \chi^B|\xi^A, \zeta^B\rangle = \langle\varphi^A|\xi^A\rangle\langle\chi^B|\zeta^B\rangle \; . \tag{7.10}$$

A basis $\{|n^A, i^B\rangle\}$ of $\mathcal{H}^{AB}$ is orthonormal if

$$\langle n^A, i^B|n'^A, i'^B\rangle = \delta_{nn'}\delta_{ii'} \tag{7.11}$$

holds, i. e. when $\{|n^A\rangle\}$ and $\{|i^B\rangle\}$ are an ONB.

The Bell basis As can readily be verified, the following four vectors make up a particular ONB in the space $\mathcal{H}^{AB} = \mathcal{H}_2^A \otimes \mathcal{H}_2^B$ of 2-qubit vectors:

$$|\Phi_\pm^{AB}\rangle := \frac{1}{\sqrt{2}}(|0^A, 0^B\rangle \pm |1^A, 1^B\rangle) \ , \quad |\Psi_\pm^{AB}\rangle := \frac{1}{\sqrt{2}}(|0^A, 1^B\rangle \pm |1^A, 0^B\rangle) \ . \tag{7.12}$$

This basis plays a special role in many investigations. We shall show later that these frequently-used *Bell states* are maximally entangled. With reference to an implementation in terms of spin polarisation states, $|\Psi_-^{AB}\rangle$ is often called a *singlet state*.

7.2.2 Operators

Product operators Let C^A be a linear operator on the space $\mathcal{H}^A$ and D^B a linear operator on $\mathcal{H}^B$. The tensor product

$$C^A \otimes D^B =: C^A D^B \tag{7.13}$$

refers to a *product operator*, which acts "space by space",

$$[C^A \otimes D^B]|\varphi^A, \chi^B\rangle = |C^A\varphi^A, D^B\chi^B\rangle \ . \tag{7.14}$$

The product operator is a linear operator on $\mathcal{H}^{AB}$

$$[C^A \otimes D^B] \sum_{n,i} \alpha_{ni}|n^A, i^B\rangle = \sum_{n,i} \alpha_{ni}|C^A n^A, D^B i^B\rangle \ . \tag{7.15}$$

The dyadic operator $|\psi^{AB}\rangle\langle\theta^{AB}|$ formed from the product vectors $|\psi^{AB}\rangle = |\varphi^A, \chi^B\rangle$ and $|\theta^{AB}\rangle = |\xi^A, \zeta^B\rangle$ is likewise a product operator

$$|\psi^{AB}\rangle\langle\theta^{AB}| = |\varphi^A, \chi^B\rangle\langle\xi^A, \zeta^B| = (|\varphi^A\rangle\langle\xi^A|) \otimes (|\chi^B\rangle\langle\zeta^B|) \ . \tag{7.16}$$

The round brackets can also be left off. The identity operator on $\mathcal{H}^{AB}$ can be dyadically expanded in terms of an ONB:

$$\mathbb{1}^{AB} = \sum_{n,i} |n^A, i^B\rangle\langle n^A, i^B| = \mathbb{1}^A \otimes \mathbb{1}^B \ . \tag{7.17}$$

With the identity operator of a factor space, product operators can be constructed which are particularly important for the physical applications. The *extended operators (subsystem operators)* which are indicated by a symbol with a hat

$$\hat{C}^A := \hat{C}^{AB} := C^A \otimes \mathbb{1}^B; \quad \hat{D}^B := \hat{D}^{AB} := \mathbb{1}^A \otimes D^B \tag{7.18}$$

are defined within $\mathcal{H}^{AB} = \mathcal{H}^A \otimes \mathcal{H}^B$, but they act only in the individual factor Hilbert spaces in a nontrivial way. They are also called *local operators*. $\hat{C}^{AB}$ and $\hat{D}^{AB}$ commute within $\mathcal{H}^{AB}$ and they obey the relations

$$\hat{C}^{AB}\hat{D}^{AB} = \hat{D}^{AB}\hat{C}^{AB} = C^A \otimes D^B \ . \tag{7.19}$$

Generalised operators Referring to the dyadic decomposition (7.17) of $\mathbb{1}^{AB}$, we can write the generalised operator Z^{AB} on $\mathcal{H}^{AB}$ in the form

$$Z^{AB} = \mathbb{1}^{AB} Z^{AB} \mathbb{1}^{AB} = \sum_{n,m} \sum_{i,j} \langle n^A, i^B | Z^{AB} | m^A, j^B \rangle (|n^A\rangle\langle m^A| \otimes |i^B\rangle\langle j^B|) \ . \quad (7.20)$$

It is determined by its matrix elements in the orthonormal basis (7.5).

The trace and partial trace The *trace* is also defined in the usual way in terms of an orthonormal basis of $\mathcal{H}^{AB}$

$$\mathrm{tr}[Z^{AB}] := \mathrm{tr}_{AB}[Z^{AB}] := \sum_{n,i} \langle n^A, i^B | Z^{AB} | n^A, i^B \rangle \ . \quad (7.21)$$

For product operators, it follows from this that

$$\mathrm{tr}_{AB}[C^A \otimes D^B] = \sum_{n,i} C^A_{nn} D^B_{ii} = \mathrm{tr}_A[C^A] \ \mathrm{tr}_B[C^B] \quad (7.22)$$

with the matrix elements C^A_{nn} and D^B_{ii}. The trace is constituted "space by space".

The computation of the *partial trace* over one of the factor spaces, for example the space $\mathcal{H}^A$, is particularly important for physical results. It is defined by

$$\mathrm{tr}_A[Z^{AB}] := \sum_n \langle n^A | Z^{AB} | n^A \rangle \ . \quad (7.23)$$

We can read off from Eq. (7.20) that an operator on $\mathcal{H}^B$ is generated in the process. For product operators, it follows that

$$\mathrm{tr}_A[C^A \otimes D^B] = \mathrm{tr}_A[C^A] D^B \ . \quad (7.24)$$

The overall trace is found to be a series of partial traces

$$\mathrm{tr}_{AB}[Z^{AB}] = \mathrm{tr}_B[\mathrm{tr}_A[Z^{AB}]] = \mathrm{tr}_A[\mathrm{tr}_B[Z^{AB}]] \ . \quad (7.25)$$

Here, the order in which the partial traces are taken is irrelevant.

The operator basis This concept also, which we have already encountered in Sect. 1.2, can be directly applied to the product space $\mathcal{H}^{AB}$. If $\{Q^A_\alpha, \ \alpha = 1, \ldots, (\dim \mathcal{H}^A)^2\}$ represents an operator basis of $\mathcal{H}^A$ and $\{R^B_\kappa, \ \kappa = 1, \ldots, (\dim \mathcal{H}^B)^2\}$ an operator basis of $\mathcal{H}^B$, then the product operators

$$T^{AB}_{\alpha\kappa} := Q^A_\alpha \otimes R^B_\kappa \quad (7.26)$$

form an orthonormal basis of the product space $\mathcal{H}^{AB}$, owing to

$$\mathrm{tr}_{AB}[T^{AB\dagger}_{\alpha\kappa} T^{AB}_{\beta\lambda}] = \delta_{\alpha\beta}\delta_{\kappa\lambda} \ . \quad (7.27)$$

Every operator Z^{AB}, which acts within $\mathcal{H}^{AB}$, can be expanded in terms of this basis:

$$Z^{AB} = \sum_{\alpha,\kappa} T^{AB}_{\alpha\kappa} \mathrm{tr}_{AB}[T^{AB\dagger}_{\alpha\kappa} Z^{AB}] \; . \tag{7.28}$$

There are operators on $\mathcal{H}^{AB}$ which cannot be written as products of two operators in the form $C^A \otimes D^B$ (cf. Sect. 7.8). But all the operators on $\mathcal{H}^{AB}$ can be written as the sum of product operators.

The product Liouville space We apply the concepts from Sect. 1.2 and form the *product Liouville space*

$$\mathbb{L}^{AB} = \mathbb{L}^A \otimes \mathbb{L}^B \; . \tag{7.29}$$

Its elements are the operators

$$C^{AB} = \sum_{\alpha,\beta} c_{\alpha\beta} Q^A_\alpha \otimes R^B_\beta \tag{7.30}$$

on $\mathcal{H}^{AB}$. The *Liouville operator* is defined through a generalisation of Eq. (1.87) with the Hamiltonian H^{AB} on $\mathcal{H}^{AB}$:

$$\mathcal{L}^{AB} Z^{AB} := \mathcal{L}^{AB}(Z^{AB}) := \frac{1}{\hbar}[H^{AB}, Z^{AB}]_- \; . \tag{7.31}$$

7.3 The Fundamentals of the Physics of Composite Quantum Systems

7.3.1 Postulates for Composite Systems and Outlook

We consider a *composite quantum system*, which itself is assumed to be isolated. Therefore, we can take over all the postulates from Chaps. 2 and 4 directly. In particular, the state of the composite system is described by a density operator in a Hilbert space. The operational interpretation of the concept "state" of a quantum system as "the system has been generated by a particular preparation procedure" holds here as well. The composite system $S^{AB\cdots}$ is supposed to consist of *subsystems* $S^A, S^B, \ldots$. *Since we wish to consider subsystems which are themselves quantum systems, it suggests itself that we associate each of them with a particular Hilbert space* $\mathcal{H}^A, \mathcal{H}^B, \ldots$ Then the only open question is what structure has the Hilbert space of the composite system, i.e. how is it composed from the $\mathcal{H}^A, \mathcal{H}^B, \ldots$. Here, there are in principle many mathematical possibilities. One is for example the direct sum $\mathcal{H}^{AB\cdots} = \mathcal{H}^A \oplus \mathcal{H}^B \oplus \ldots$. However, one in fact postulates the tensor product as described in Sect. 7.2.1, in order to obtain agreement with experiments. This specification has far-reaching consequences for all physical statements about composite quantum systems. We shall be interested in precisely these statements in the following sections.

The postulate *The states of an isolated composite system $S^{AB\cdots}$ which is composed of the subsystems $S^A, S^B, \ldots$ are described by density operators $\rho^{AB\cdots}$ in the product Hilbert space*

$$\mathcal{H}^{AB\cdots} = \mathcal{H}^A \otimes \mathcal{H}^B \otimes \ldots \tag{7.32}$$

The postulates for isolated systems from Sect. 2.1 and Sect. 4.2 can be applied to the overall system $S^{AB\cdots}$. If a system is not isolated, it can be made into an isolated system by including the "rest of the world". It then becomes itself a subsystem.

Outlook We can immediately read off a series of special properties of composite systems from this postulate. The mathematical product structure (7.32) defines an organisation scheme. We demonstrate it using the example of a bipartite system S^{AB}.

(i) States: a pure state can be a product state $|\psi^{AB}\rangle = |\phi^A\rangle \otimes |\chi^B\rangle$ or an entangled state $|\psi^{AB}\rangle \neq |\phi^A\rangle \otimes |\chi^B\rangle$ (compare Sect. 7.2.1). The unusual properties of entangled states, in particular the appearance of non-classical correlations and their applications, will be discussed in the rest of this chapter and in all the remaining chapters in detail. We consider correlated density operators $\rho^{AB} \neq \rho^A \otimes \rho^B$ in Sect. 8.1.

(ii) Observables: there is a special case of the extended observable operators, such as $\hat{C}^{AB} = C^A \otimes \mathbb{1}^B$, which is generated from an observable operator which acts on only one of the product spaces. These describe *local measurements* which are carried out on only one of the subsystems (e. g. a measurement of the observable C^A on the subsystem S^A). There are however more general Hermitian operators on $\mathcal{H}^{AB}$ (e.g. $Z^{AB} = C^A \otimes D^B + E^A \otimes F^B$), which cannot be expressed as extended operators. They also correspond to projective measurements of physical observables Z^{AB}. These latter observables are called *non-local observables* or *collective observables*. The corresponding measurements are *non-local measurements*, which in general cannot be carried out directly as local measurements on S^A and S^B. This holds also for the special case of the observables which correspond mathematically to operator products (e. g. $Z^{AB} = C^A \otimes D^B$), but cannot be implemented physically as local measurements of the extended observables ($C^A \otimes \mathbb{1}^B$ and $\mathbb{1}^A \otimes D^B$). Non-local measurements are important in connection with quantum correlations and non-local information storage. We will therefore discuss them only in Sect. 9.2.

(iii) Unitary evolution: the unitary evolutions also need not have the structure $U^{AB} = U^A \otimes U^B$. There can be for example an interaction between the systems S^A and S^B. We discuss this in Sect. 7.6. Non-local unitary evolution can act to entangle and to disentangle states. In order for a composite system to be in an entangled state, dynamic interactions between the subsystems must not exist at the same time.

(iv) The postulate (7.32) provides the required possibility of separate interventions and therefore the resolution of the composite system into subsystems. Not only local observable operators, but rather all local operators which act on a subsystem commute with all the local operators which act on some other subsystem (cf. Eq. (7.19)). This does not depend on the order in which the corresponding actions occur. Thus, in measurements on

subsystems, the correlations between the measured values obtained become an important quantity. They are characterised by the joint probabilities for the occurrence of the measured values.

7.3.2 The State of a Subsystem, the Reduced Density Operator, and General Mixtures

Via the postulate, the details of the projective measurement of an observable of the composite systems are determined. This measurement on the composite system is described by an Hermitian operator on $\mathcal{H}^{AB\cdots}$. The measurement of an observable on a subsystem, e. g. on S^A, is included as a special case. It is associated with an observable operator C^A which acts on $\mathcal{H}^A$. This *local measurement* corresponds in $\mathcal{H}^{AB\cdots}$ to a *local observable*

$$\hat{C}^{AB\ldots E} = C^A \otimes \mathbb{1}^B \otimes \ldots \otimes \mathbb{1}^E \; . \tag{7.33}$$

In this chapter, we shall restrict ourselves to composite systems which are composed of two subsystems. The extension to a greater number of subsystems is straightforward.

Probability statements According to the postulate, the rules for the measurement dynamics apply also to the states ρ^{AB} of the composite system S^{AB}. We investigate the resulting consequences for local measurements. To this end, it is expedient to associate to each subsystem a *reduced density operator* by taking the partial trace over the other subsystem

$$\rho^A := \mathrm{tr}_B \left[\rho^{AB}\right], \quad \rho^B := \mathrm{tr}_A \left[\rho^{AB}\right] \; . \tag{7.34}$$

Since ρ^{AB} is a density operator, ρ^A and ρ^B likewise fulfill the conditions for being density operators. The eigenvalue equation of the observable C^A,

$$C^A |c_n^{(r)A}\rangle = c_n |c_n^{(r)A}\rangle, \quad r = 1, \ldots, g_n \tag{7.35}$$

leads to the ONB $\{|c_n^{(r)A}\rangle\}$ of $\mathcal{H}^A$ and the eigenvalues $\{c_n\}$ with the degeneracies g_n. The probability of obtaining the measured value c_n from a measurement of C on the system S^A is then given by the *local projection operator*

$$\hat{P}_n^A = P_n^A \otimes \mathbb{1}^B, \quad P_n^A := \sum_{r=1}^{g_n} |c_n^{(r)A}\rangle\langle c_n^{(r)A}| \tag{7.36}$$

through the mean value

$$p(c_n) = \mathrm{tr}_{AB}[\hat{P}_n^A \rho^{AB}] = \mathrm{tr}_A[\mathrm{tr}_B\{\hat{P}_n^A \rho^{AB}\}] = \mathrm{tr}_A[P_n^A \rho^A] \; . \tag{7.37}$$

In a similar manner, for the expectation value of the observables C, we obtain

$$\langle \hat{C}^A \rangle = \mathrm{tr}_{AB}[\rho^{AB} \hat{C}^A] = \mathrm{tr}_A[\rho^A C^A] \; . \tag{7.38}$$

To summarise, we can conclude that: *all probability statements about local measurements on a subsystem S^A are obtained by associating the reduced density operator ρ^A from Eq. (7.34) to the system S^A and applying the rules postulated for the density operators of isolated systems.*

The state of a subsystem Since all probability statements for measurements on S^A are unambiguously determined by the reduced density operator ρ^A, it is tempting to say that the subsystem S^A *is in the state* ρ^A. Thus, in Chap. 2, we introduced the concept of a state. The composite system S^{AB} passes through a preparation procedure which leads to the state ρ^{AB}. Together with it, the state $\rho^A = \mathrm{tr}_B\left[\rho^{AB}\right]$ is prepared.

General mixtures If the composite system S^{AB} is in the product state $|\alpha_k^A, \beta_k^B\rangle$, then the subsystem S^A is in the pure state $|\alpha_k^A\rangle$. If the state of S^{AB} is, in particular, a statistical mixture (blend or *proper mixture*) of such product states (cf. Chap. 4),

$$\rho^{AB} = \sum_s p_s |\alpha_s^A, \beta_s^B\rangle\langle\alpha_s^A, \beta_s^B| = \sum_s p_s |\alpha_s^A\rangle\langle\alpha_s^A| \otimes |\beta_s^B\rangle\langle\beta_s^B|, \quad \sum_s p_s = 1 \ , \tag{7.39}$$

then S^A or S^B are likewise statistical mixtures

$$\rho^A = \mathrm{tr}_B[\rho^{AB}] = \sum_s p_s |\alpha_s^A\rangle\langle\alpha_s^A| \, , \quad \rho^B = \sum_s p_s |\beta_s^B\rangle\langle\beta_s^B| \tag{7.40}$$

of the states $|\alpha_k^A\rangle$ or $|\beta_S^B\rangle$. They were produced by the preparation procedure and are present as real states. An ignorance interpretation (compare Sect. 4.3) is possible. Equation (7.40) is obtained from (7.24) and the dyadic decomposition of $\mathbb{1}^A$ or $\mathbb{1}^B$.

In general, the state of a quantum systems S^{AB} will not be a statistical mixture as in Eq. (7.39). The state of the subsystem S^A is then also not a statistical mixture. An ignorance interpretation is not possible. Nevertheless, the state is described by a reduced density operator ρ^A. We therefore employ the concept *mixture* to this state ρ^A of S^A also, although – as already mentioned in Sect. 4.2 – no "mixing" has occurred; and we simply leave off the adjective "statistical" for clarity. In this case, one also speaks of the state as an *improper mixture* in contrast to a *proper mixture*. "Mixture" is thus the umbrella term.

To make this clear, we can consider for example a system S^{AB} which is in a Bell state. In this case, the states of the subsystems are maximally mixed as a result of the entanglement

$$\rho^A = \mathrm{tr}_B[\Phi_\pm^{AB}] = \frac{1}{2}\mathbb{1}^A \, , \quad \rho^A = \mathrm{tr}_B[\Psi_\pm^{AB}] = \frac{1}{2}\mathbb{1}^A \, . \tag{7.41}$$

A corresponding relation holds for S^B. However, S^{AB} was prepared in a pure state.

In quantum systems, the states of subsystems can be mixtures which – with respect to their preparation – are not statistical mixtures and therefore do not permit an ignorance interpretation. For their density operators, there are formally arbitrarily many ensemble decompositions. There are therefore arbitrarily many statistical mixtures of an isolated individual system, with which they can be indistinguishably *simulated* with respect to all probability statements for local measurements. By means of local measurements, one cannot determine whether a density operator ρ^A belongs to an individual system S^A or is rather a reduced density operator of a subsystem S^A which is part of a larger system. This again justifies the application of the term "mixture" to all reduced density operators. We mention finally that mixtures in classical physics are always statistical mixtures. We shall return to the connection with entanglement later in Sect. 8.1.

7.4 Manipulations on a Subsystem

7.4.1 Relative States and Local Unitary Transformations

Relative states Making use of the ONB $\{|c_n^A\rangle\}$ and $\{|d_i^B\rangle\}$ of $\mathcal{H}^A$ or $\mathcal{H}^B$, we can write the pure state $|\psi^{AB}\rangle$ of the composite system S^{AB} as a decomposition

$$|\psi^{AB}\rangle = \sum_{n,i} \alpha_{ni} |c_n^A, d_i^B\rangle \tag{7.42}$$

in terms of the basis vectors. It proves expedient to split up the double sum in the form

$$|\psi^{AB}\rangle = \sum_n |c_n^A, \tilde{w}_n^B\rangle \tag{7.43}$$

with

$$|\tilde{w}_n^B\rangle := \sum_i \alpha_{ni} |d_i^B\rangle; \quad |w_n^B\rangle = \frac{|\tilde{w}_n^B\rangle}{\sqrt{\langle \tilde{w}_n^B | \tilde{w}_n^B\rangle}} . \tag{7.44}$$

The vector $|w_n^B\rangle$ describes the *relative state* belonging to $|c_n^A\rangle$. Non-normalised states are again denoted by a tilde. The relative vectors $|w_n^B\rangle$ in general do not make up an orthonormal system. Their number need not be the same as the dimension of the Hilbert state $\mathcal{H}^B$. $|\psi^{AB}\rangle$ can, in analogy to Eq. (7.43), also be decomposed in terms of the relative states $|\tilde{v}_i^A\rangle$ belonging to the $|d_i^B\rangle$:

$$|\psi^{AB}\rangle = \sum_i |\tilde{v}_i^A, d_i^B\rangle . \tag{7.45}$$

Local unitary manipulations We now allow a unitary dynamics to act upon the subsystem S^A

$$\hat{U}^{AB} = U^A \otimes \mathbb{1} . \tag{7.46}$$

It produces the transition

$$|\psi^{AB}\rangle \to |\psi'^{AB}\rangle = \sum_n |U^A c_n^A\rangle |\tilde{w}_n^B\rangle . \tag{7.47}$$

Here, in general the state of S^A is changed, and in particular that of S^{AB}. The vectors $|U^A c_n^A\rangle$ again represent an ONB of $\mathcal{H}^A$; thus, for the state of S^B we have the unchanged result

$$\rho^B \to \rho'^B = \rho^B = \sum_n |\tilde{w}_n^B\rangle\langle \tilde{w}_n^B| . \tag{7.48}$$

We take as an example of this a transition between two vectors of the Bell basis (cf. Eq. (7.12)) to which we shall return later:

$$(\sigma_1^A \otimes \mathbb{1}^B) |\Psi_+^{AB}\rangle = |\Phi_+^{AB}\rangle . \tag{7.49}$$

In this special case, not only the reduced density operator of S^B remains unchanged, but also that of S^A: $\rho'^A = \rho'^B = \rho^A = \rho^B = \frac{1}{2}\mathbb{1}$.

A dynamic manipulation which effects a unitary transformation of the subsystem S^A has no influence on the state of the other subsystem S^B (and vice versa*). Even when an entangled state is present, Bob can by no means determine via measurements on his subsystem S^B whether Alice has carried out a unitary manipulation on her subsystem.* One can readily convince oneself that this statement is still true if the state of S^{AB} is a mixture.

7.4.2 Selective Local Measurements

The resulting state of the composite systems We again consider a quantum system S^{AB} which is composed of the (sub) systems S^A and S^B. We wish to measure the observable C on the subsystem S^A and the observable D on the subsystem S^B (*local measurements*). The associated observable operators $\hat{C}^A = C^A \otimes \mathbb{1}^B$ and $\hat{D}^B = \mathbb{1}^A \otimes D^B$ commute

$$[\hat{C}^A, \hat{D}^B]_- = 0\,. \tag{7.50}$$

We note also the corresponding eigenvalue equations

$$C^A|c_n^A\rangle = c_n|c_n^A\rangle, \quad D^B|d_i^B\rangle = d_i|d_i^B\rangle\,. \tag{7.51}$$

The vectors $|c_n^A\rangle$ and $|d_i^B\rangle$ make up an ONB of $\mathcal{H}^A$ or $\mathcal{H}^B$. The possible measured values c_n and d_i resulting from the local measurements are assumed for simplicity not to be degenerate.

We first carry out measurements only on the subsystem S^A and apply the postulate from Sect. 7.3.1. A measurement of the observable C on the subsystem S^A, in which a selection among the results c_n of the measurements is made, transforms the state ρ^{AB} of the composite system S^{AB} into the (non-normalised) state $\tilde{\rho}'^{AB}$

$$\rho^{AB} \to \tilde{\rho}'^{AB}_n = \hat{P}^A_n \rho^{AB} \hat{P}^A_n\,. \tag{7.52}$$

The projection operator $\hat{P}^A_n$ is defined in Eq. (7.36). One can read off from Eq. (7.37) that the trace of the resulting non-normalised density operator $\tilde{\rho}'^{AB}_n$ again gives directly the probability $p(n)$ that a measurement will lead to the value c_n (cf. Eq. (4.23))

$$p(c_n) = \mathrm{tr}[\tilde{\rho}'^{AB}_n]\,. \tag{7.53}$$

If the composite system S^{AB} was previously in the pure state $|\psi^{AB}\rangle$ of Eq. (7.42), then the selective measurement causes the transition

$$|\psi^{AB}\rangle \to |\tilde{\psi}'^{AB}_n\rangle = \hat{P}^A_n|\psi^{AB}\rangle = |c_n^A\rangle \otimes \sum_i \alpha_{ni}|d_i^B\rangle = |c_n^A\rangle \otimes |\tilde{w}_n^B\rangle\,. \tag{7.54}$$

The subsystem S^B transforms into the relative state $|w_n^B\rangle$ of Eq. (7.44). For an entangled state $|\psi^{AB}\rangle$, a non-degenerate selective measurement on a subsystem breaks the entanglement.

Furthermore, we find as a special case of Eq. (7.53): *the probability of obtaining the measured value c_n is, from Eq. (7.37), given by the square of the norm of the non-normalised relative state vector $|\tilde{w}_n^B\rangle$:*

$$p(c_n) = \langle\psi^{AB}|\left(|c_n^A\rangle\langle c_n^A| \otimes \mathbb{1}^B\right)|\psi^{AB}\rangle = \langle\tilde{w}_n^B|\tilde{w}_n^B\rangle = ||\tilde{w}_n^B||^2\,. \tag{7.55}$$

$p(c_n)$ can also be written as a function of the expansion coefficients α_{ni} of Eq. 7.42:

$$p(c_n) = \sum_i |\alpha_{ni}|^2 \,. \tag{7.56}$$

The resulting state of the subsystem The reduced density operator ρ^A of the subsystem S^A is transformed into $\tilde{\rho}_n'^A$ by a selective measurement:

$$\rho^A \to \tilde{\rho}_n'^A = \mathrm{tr}_B[\tilde{\rho}_n'^{AB}] = \mathrm{tr}_B[\hat{P}_n^A \rho^{AB} \hat{P}_n^A] \,. \tag{7.57}$$

Insertion of $\hat{P}_n^A$ and normalisation leads with Eqs. (7.36) and (7.37) to

$$\rho^A \to {\rho'_n}^A = \frac{P_n^A \rho^A P_n^A}{\mathrm{tr}_A[P_n^A \rho^A]} = \frac{P_n^A \rho^A P_n^A}{p(c_n)} \,. \tag{7.58}$$

If the initial state is the pure state $|\psi^{AB}\rangle$, then we obtain

$$\rho_n'^A = |c_n^A\rangle\langle c_n^A| \,. \tag{7.59}$$

S^A is in the state $|c_n^A\rangle$ after the measurement. This also follows directly from Eq. (7.54).

Operational description It is helpful to make it clear on an operational level just how a selective measurement of Eq. (7.54) is carried out in practice and how the state $|\psi'^{AB}\rangle$ (cf. Eq. (7.54)) is produced. As we have seen in Sect. 2.1.2, state vectors are associated with preparation procedures. How is the corresponding preparation procedure for $|\psi'^{AB}\rangle$ carried out? Many individual bipartite systems have passed through the preparation device for $|\psi^{AB}\rangle$. The single system S^{AB} can for example consist of a photon moving to the left and one moving to the right in the state $|\psi^{AB}\rangle$. Alice measures (on the subsystem S^A, left photon) the observable C without annihilating the system. Those complete bipartite systems (photon pairs) from which Alice has obtained the measured value c_n are sorted out. Only they are used for further manipulations. This is the significance of Eq. (7.54). All the remaining bipartite systems are eliminated and no longer take part in future experiments.

To ensure that in fact complete bipartite systems (photon pairs) are sorted out, Bob must also act and eliminate his subsystem (right photon) when Alice has eliminated hers. In the example of the photons, he cannot simply let them all continue on their way. In order that he allow the correct ones to continue, Alice must give him the information for each photon pair as to whether she has sorted out her photon or not. Those photon pairs which then finally are allowed to continue are all in the state $|\psi'^{AB}\rangle = |c_n^A, w_n^B\rangle$. The overall procedure, which also includes an exchange of information, then prepares the subsystem S^B (photon at Bob's location) in the state $|w_n^B\rangle$. Bob can also number his photons, store them and later, following Alice's instructions, he can sort them. *A selective local measurement is a preparation procedure for the overall system, which is based on a selective measurement on a subsystem and on classical communication. It requires a selection process for both subsystems.*

7.4.3 A Non-Selective Local Measurement

The resulting state of the composite system Following a measurement of the observable C on the system S^A, in which no selection according to the measured values is carried out (cf. Fig. 7.1), for the composite system S^{AB}, the state $\tilde{\rho}'^{AB}_{n.s.}$ is present:

$$\rho^{AB} \xrightarrow{n.s.} \rho'^{AB}_{n.s.} = \sum_n p(c_n) \frac{\tilde{\rho}'^{AB}_n}{\mathrm{tr}[\tilde{\rho}'^{AB}_n]} = \sum_n \tilde{\rho}'^{AB}_n \ . \tag{7.60}$$

This follows immediately from equations (7.52) and (7.53). For the pure initial state $|\psi^{AB}\rangle$, we obtain, corresponding to Eq. (7.54)[1]:

$$|\psi^{AB}\rangle \xrightarrow{n.s.} \rho'^{AB}_{n.s.} = \sum_n |c^A_n, \tilde{w}^B_n\rangle\langle c^A_n, \tilde{w}^B_n| = \sum_n |c^A_n\rangle\langle c^A_n| \otimes |\tilde{w}^B_n\rangle\langle \tilde{w}^B_n| \ . \tag{7.61}$$

The superposition of Eq. (7.42) has been decomposed into the mixture of Eq. (7.61).

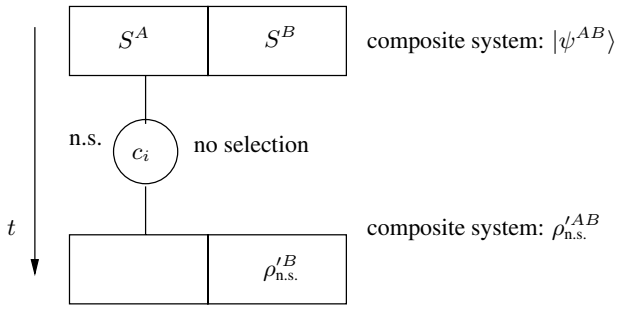


Figure 7.1: A non-selective measurement on the subsystem S^A.

The resulting states of the subsystems The state of the subsystem S^A after the non-selective measurement is given by the reduced density operator. With Eqs. (7.60) and (7.52), we obtain

$$\rho^A \xrightarrow{n.s.} \rho'^A_{n.s.} = \mathrm{tr}_B[\rho'^{AB}_{n.s.}] = \mathrm{tr}_B[\sum_n \hat{P}^A_n \rho^{AB} \hat{P}^A_n] \ . \tag{7.62}$$

Carrying out the trace with $\hat{P}^A_n = P^A_n \otimes \mathbb{1}^B$ leads to

$$\rho^A \xrightarrow{n.s.} \rho'^A_{n.s.} = \sum_n P^A_n \rho^A P^A_n \ . \tag{7.63}$$

As we saw in Sect. 7.3.2, the state of a subsystem is represented by the corresponding reduced density operator. Probability statements are obtained by following the rules for density operators in Chap. 4. The comparison of Eq. (7.58) with Eq. (4.19) and Eq. (7.63) with

[1] We shall see in Sect. 8.1 that the resulting state is not entangled.

(4.25) shows that: *for the transitions between the reduced density operators as produced by selective or non-selective local measurements on a subsystem, the rules for density operators from Chap. 4 can be applied.*

What can we say about the other subsystem S^B? All the measurements on Bob's subsystem S^B in the case of non-selective measurements by Alice on S^A can be described by the reduced density operator

$$\rho'^{B}_{n.s.} = \mathrm{tr}_A[\rho'^{AB}_{n.s.}] \; . \tag{7.64}$$

We reformulate it with the aid of Eqs. (7.60) and (7.52) and find by using $\sum_n \hat{P}^A_n = \mathbb{1}^{AB}$ the result:

$$\rho'^{B}_{n.s} = \mathrm{tr}_A[\sum_n \tilde{\rho}'^{AB}_n] = \mathrm{tr}_A[\sum_n \hat{P}^A_n \rho^{AB}] = \mathrm{tr}_A[(\sum_n \hat{P}^A_n)\rho^{AB}] = \mathrm{tr}_A \rho^{AB} = \rho^B \; . \tag{7.65}$$

The density operator $\rho'^{B}_{n.s}$ of the subsystem S^B after the non-selective measurement on S^A is the same as the density operator ρ^B before the measurement.

This is a remarkable result. Let us consider the situation in which the system S^A is at Alice's location and the system S^B at Bob's (spatially separated) location. In a preparation procedure, bipartite systems are often produced in the state ρ^{AB}. It can be entangled. It is then left open to Alice as to whether she carries our measurements of some observable C on her system or not. *Bob cannot determine in any manner by measurements on his subsystem S^B whether or not Alice has carried out measurements.* The analogous statement for unitary manipulations on the system S^A has already been derived in Sect. 7.4.1.

7.5 Separate Manipulations on both Subsystems

7.5.1 Pairs of Selective Measurements

First, Alice carries out a measurement and obtains the result c_n with the probability $p(c_n) = \langle \tilde{w}^B_n | \tilde{w}^B_n \rangle$. The system is transformed after the selection described above into the composite state $|c^A_n, w^B_n\rangle$ (compare Fig. 7.2). If Bob makes a measurement following this selection, he obtains the value d_i with the conditional probability

$$p(d_i|c_n) = \frac{|\alpha_{ni}|^2}{p(c_n)} \; . \tag{7.66}$$

This can be read off from Eqs. (7.44) and (7.55). The composite system is then transformed into the product state $|c^A_n, d^B_i\rangle$ after another selection for which Bob informs Alice of the result he has obtained. If, in reverse, first Bob and then – after selection according to the measured value d_i – Alice makes a measurement, we obtain analogously (see Fig. 7.2) after the second selection the same final state for the pair of measured values (c_n, d_i). For the probabilities, we then have

$$p(c_n|d_i) = \frac{|\alpha_{ni}|^2}{p(d_i)} \; . \tag{7.67}$$

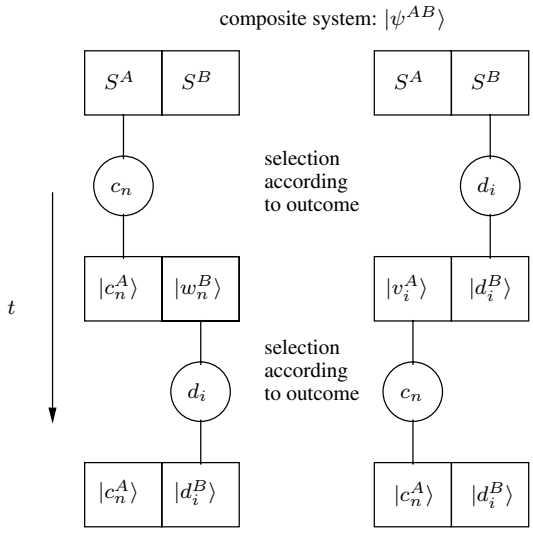


Figure 7.2: A selective measurement on the subsystems S^A and S^B. *Left*: the measurement is first carried out on S^A and then on S^B; *right* in the reverse order. A selection is carried out in each case according to the measured values d_i and c_n. The probability of obtaining the pair of measured values (c_n, d_i) and the corresponding final state $|c_n^A, d_i^B\rangle$ is the same in both cases.

The joint probability $p(c_n, d_i)$ with which the pair of measured values (c_n, d_i) is obtained from selective local measurements is independent of the order in which they are carried out. One finds

$$p(c_n, d_i) = p(c_n|d_i)p(d_i) = p(d_i|c_n)p(c_n) = |\alpha_{ni}|^2 = \langle\psi^{AB}|P_{ni}^{AB}|\psi^{AB}\rangle \tag{7.68}$$

with the projection operator

$$P_{ni}^{AB} := |c_n^A, d_i^B\rangle\langle c_n^A, d_i^B|\,. \tag{7.69}$$

The final state is $P_{ni}^{AB}|\psi^{AB}\rangle = |c_n^A, d_i^B\rangle$. Since the observables $\hat{C}^A$ and $\hat{D}^B$ of the local measurements commute, this could have been expected. We add that all of the statements made above for the pure initial state $|\psi^{AB}\rangle$ can be applied in the well-known manner when the initial state is a mixture with the density operator ρ^{AB}.

Instead of selecting after each local measurement, Alice and Bob can also find the state $|c_n, d_i\rangle$ after a large number of measurements by referring to the result (c_n, d_i). In this case, also, an exchange of information in both directions is necessary. It is a part of the preparation procedure for the state $|c_n^A, d_i^B\rangle$. In many cases, one is interested in the probabilities $p(c_n, d_i)$ with which the pairs of measured values (c_n, d_i) occur. To find them, Alice and Bob meet after carrying out measurements on many systems and determine the relative frequency of the combinations of measured values. These *correlations* of locally-obtained results are determined by the preparation procedure of the initial state (7.42).

Mean values The dyadic decomposition of the operators $C^A \otimes D^B$ has the form (compare Eq. (7.51))

$$C^A \otimes D^B = \sum_{n,i} c_n d_i |c_n^A, d_i^B\rangle\langle c_n^A, d_i^B| \,. \tag{7.70}$$

For its mean value in the state ρ^{AB}, we have

$$\sum_{n,i} \mathrm{tr}_{AB}\left[P_{ni}^{AB}\rho^{AB}\right] c_n d_i = \mathrm{tr}_{AB}\left[(C^A \otimes D^B)\,\rho^{AB}\right] \,. \tag{7.71}$$

The trace on the left side is the probability that in local measurements of $\hat{C}^A$ and $\hat{D}^B$ on the subsystems S^A and S^B, the pair of measured values (c_n, d_i) will be obtained. *The mean value of the products of correlated local measured values is the same as the mean value of the product operator.* We will make use of this fact, especially in Sects. 9.2.2 and 10.1, in connection with non-local measurements.

7.5.2 Non-Local Effects: "Spooky Action at a Distance"?

For an improved understanding, it is helpful to confront the results of the preceding sections with a popular catchword. We consider the following situation: we carry out local measurements of the same observable C on a system in the state

$$|\psi^{AB}\rangle = \frac{1}{\sqrt{2}}(|c_1^A, c_1^B\rangle + |c_2^A, c_2^B\rangle) \tag{7.72}$$

(conventions as in Sect. 7.4.2). The possible results are c_1 or c_2. The probabilities of occurrence of the pairs of measured values are

$$p(c_1, c_1) = p(c_2, c_2) = \frac{1}{2} \tag{7.73}$$

$$p(c_1, c_2) = p(c_2, c_1) = 0 \,. \tag{7.74}$$

The measurements on S^A yield, for example, the value c_1. Then one often says, in an abbreviated and sometimes misleading manner of speech, that the measurement has transformed the composite system S^{AB} into the state $|c_1^A, c_2^B\rangle$ and thereby the subsystem S^B into the state $|c_1^B\rangle$. This holds independently of the spatial separation between the system S^A at Alice's location and S^B at Bob's. In popular-scientific descriptions, this is often referred to as "spooky action at a distance"[2]. Is the situation of quantum physics correctly characterised by this term?

We have seen the the preparation of quantum objects in a state $|c_1^A, c_1^B\rangle$ requires a selection by Alice, a communication at most at the velocity of light between Alice and Bob, and a selection by Bob. This is most certainly not a case of instantaneous action at a distance.

[2] A. Einstein wrote concerning the quantum theory: "Ich kann aber deshalb nicht ernsthaft daran glauben, weil die Theorie mit dem Grundsatz unvereinbar ist, dass die Physik eine Wirklichkeit in Zeit und Raum darstellen soll, ohne spukhafte Fernwirkung". (A. Einstein in a letter to M. Born dated 3.3.1947 [EB 69]). Born's translation: "I cannot seriously believe in it because the theory cannot be reconciled with the idea that physics should represent a reality in time and space, free from spooky actions at a distance". We shall return to what Einstein understood to be "reality" in Chap. 10.

Perhaps the catchword is not meant to refer to states, and thereby to preparation procedures, but rather to measured values. If Alice obtains the value c_1, then Bob, according to Eq. (7.74), will with certainty find the value c_1. This is also the case when the two measurements are carried out simultaneously. For a system prepared in the state (7.72), measurements of the observable C on S^A and S^B yield completely correlated results. However, the occurrence of the two results is not causally related. We are familiar with a similar situation in the case of classical systems: as a preparation procedure, either a red or a blue ball is placed into each of two boxes. If the preparation procedure is known, then after opening one of the boxes, it can be predicted with certainty what the result of a simultaneous observation of the ball in the other box will be. It is not necessary that the colour of the one ball be somehow connected with the colour of the other via some interaction which propagates with more than the velocity of light. *Correlations are already determined by the preparation procedure*.

By the comparison to the two-ball experiment, we wished to emphasize that in this situation, the correlations are decisive. Not all statements about composite quantum systems can be simulated by classical systems such as e.g. coloured balls. We will discuss this in detail in Chap. 10. The section after the next contains a first demonstration of this.

How Alice prepares Bob's subsystem in a state of her choice For every given ONB $\{|s\rangle, |t\rangle\}$, the state $|\Phi_+^{AB}\rangle = \frac{1}{\sqrt{2}}(|0^A, 1^B\rangle - |1^A, 2^B\rangle)$ can always be written in the form

$$|\Phi_+^{AB}\rangle = \frac{1}{\sqrt{2}}(|s^A, t^B\rangle - |t^A, s^B\rangle) \; . \tag{7.75}$$

Alice wants to transform the subsystem S^B at Bob's location into the state $|s^B\rangle$. To do so, she makes a measurement on her system S^A in the ONB $\{|s^A\rangle|t^A\rangle\}$ and informs Bob if her measurement yields the result associated with $|t^A\rangle$. Then Bob can select accordingly among his subsystems. The result is a preparation procedure which leads Bob to quantum objects in the state $|s^B\rangle$. For this purpose, no quantum objects need be transmitted between Alice and Bob. The entangled state serves as a tool (similar procedures are described in Sect. 11.2 and in connection with quantum teleportation in Sect. 11.3).

7.6 The Unitary Dynamics of Composite Systems

We consider unitary transformations of the composite system. The von Neumann equation (4.9) or (4.10) can be applied to composite systems according to the postulates

$$i\hbar\frac{\mathrm{d}\rho^{AB}}{\mathrm{d}t} = [H^{AB}, \rho^{AB}(t)]_- \qquad i\frac{\mathrm{d}\rho^{AB}}{\mathrm{d}t} = \mathcal{L}^{AB}\rho^{AB}(t) \; . \tag{7.76}$$

with the Liouville operator $\mathcal{L}^{AB} \in \mathbb{L}^A \otimes \mathbb{L}^B$. We employ the Schrödinger representation. If an interaction described by the Hamiltonian $H_{\text{int}}^{AB} \neq 0$ is present between the subsystems S^A and S^B, then the individual subsystems are *open* quantum systems. The overall Hamiltonian then has the form

$$H^{AB} = H^A \otimes \mathbb{1}^B + \mathbb{1}^A \otimes H^B + H_{\text{int}}^{AB} \; . \tag{7.77}$$

The associated Liouville operator is found to be

$$\mathcal{L}^{AB} = \mathcal{L}^A + \mathcal{L}^B + \mathcal{L}^{AB}_{\text{int}} \tag{7.78}$$

and it follows for the von Neumann equation:

$$i\frac{\mathrm{d}\rho^{AB}}{\mathrm{d}t} = (\mathcal{L}^A + \mathcal{L}^B + \mathcal{L}^{AB}_{\text{int}})\rho^{AB}(t) \; . \tag{7.79}$$

This leads to a differential equation for the reduced density operator ρ^A

$$i\frac{\mathrm{d}\rho^{A}}{\mathrm{d}t} = \mathcal{L}^A\rho^A(t) + \mathrm{tr}_B[\mathcal{L}^{AB}_{\text{int}}\rho^{AB}(t)] \; . \tag{7.80}$$

To determine $\rho^A(t)$, the complete equation (7.79) must be integrated. There are various approximation methods to accomplish this. In Sect. 13.1 and Chap. 14, we will encounter an in-out approach to the dynamics of open systems, which is not based upon the differential time dependence of $\rho^A(t)$ described by Eq. (7.80). Instead, it relates the final state $\rho^A(t_{\text{out}})$ to the initial state $\rho^A(t_{\text{in}})$ by means of a superoperator.

7.7 A First Application of Entanglement: a Conjuring Trick

In the coming chapters, we will demonstrate repeatedly that entanglement is a central tool on which the effects of quantum information theory are based. Entanglement and the quantum correlations which arise from it can however also be a tool for the study of the fundamentals of quantum theory. We wish to demonstrate this in answering the following underlying question: can effects of quantum theory be explained by means of classical physics – possibly in the framework of theories which have yet to be formulated? This will give us directly an example of an application for the formalism introduced in the preceding sections. In a wider context, we will come back to this question in Chap. 10.

7.7.1 The Conjuring Trick

A magician amazes his audience with the following trick: the audience sees the magician give something to his two assistants Alice and Bob. Alice and Bob then each go into separate rooms which are perfectly insulated against any exchange of information. In each room is an audience. In each, a coin is tossed and, depending on the result of the toss, a question is asked of Alice or Bob. If the result is “heads”, then the question concerns the favourite colour; it can be answered with either “red” or “green”. If the tossed coin gives “tails”, then the audience is to ask the question, “What is your favourite vegetable”, and the answer can be either “carrots” or “peas”. The question and answer are written down; one round is then finished. Alice, Bob and the magician meet again, enter the question-and-answer pair in a list with the audience as witnesses, and repeat the whole procedure again from the beginning. A large number of such rounds is completed. At the end, the combined audience analyses the question-and-answer pairs, looking for correlations.

There are four pairs of questions which can be divided into three cases: both are asked for a colour, one for a colour and the other for a vegetable, or both are asked for vegetables. In each performance, the following correlations are found:

	Alice	**Bob**
1st Case	*colour?*	*colour?*
	green!	green!

To the pair of questions (colour?, colour?), the pair of answers (green!, green!) is given with a non-vanishing frequency.

2nd Case	*colour?*	*vegetable?*
	green!	peas!
	vegetable?	*colour?*
	peas!	green!

When one answers this combination of questions with "green!", then the other always gives the appropriate answer "peas".

3rd Case	*vegetable?*	*vegetable?*

To this combination of questions, one of the two assistants with certainty gives the answer "carrots!".

This is what the audience records.

7.7.2 Classical Correlations can give no Explanation

The audience sees that pairs of answers are given with a certain regularity. How did the magician arrange this? What was his trick? The audience presumes that Alice and Bob were given slips of paper on which a colour and a vegetable were written. They then read off the answers to the questions correspondingly. The magician had prepared pairs of paper slips with different abundances, so that precisely the correlations between the answers as listed above would be observed.

Assuming this were the case, then to produce the first case in the table, the magician must have given out slips which both had the colour "green" written on them. In order that the pairs of questions (colour?, vegetable?) and (vegetable?, colour?) would always be answered correctly, the vegetable on both of these slips would have to be "peas" (second case). However, this would contradict the requirement that in the case of the third possible combination of questions (vegetable?, vegetable?) at least one of the assistants would answer with "carrots". The method with slips of paper thus does not yield the results observed. We want to demonstrate that the magician nevertheless need not possess paranormal abilities in

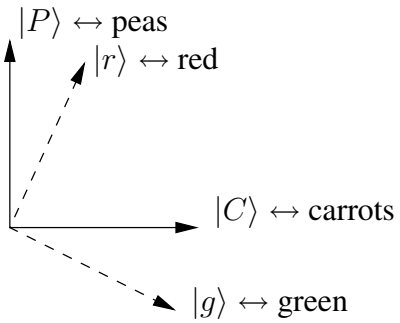


Figure 7.3: Bases in which the measurements are carried out for the conjuring trick. When pairs of photons are used, these are the analyser devices.

order to cause the observed correlations of the answers. It suffices for him to have some knowledge of entangled states.

7.7.3 The Trick

The magician's trick consists of the fact that he does not use correlated classical systems such as pairs of paper slips, but instead he makes use of entangled quantum systems. He gave to Alice and Bob each a subsystem of a bipartite systems, which was prepared to be in the state

$$|\chi^{AB}\rangle = N(|r^A,\, r^B\rangle - a^2|P^A,\, P^B\rangle) \tag{7.81}$$

with $a \in \mathbb{R}$ and $a \neq 0, a \neq 1$. N is a normalisation factor and $\{|r\rangle, |g\rangle\}$ and $\{|C\rangle, |P\rangle\}$ are orthonormal bases of $\mathcal{H}^2$, which are rotated relative to one another (see Fig. 7.3).

$$|r\rangle = a|P\rangle + b|C\rangle \tag{7.82}$$
$$|g\rangle = b|P\rangle - a|C\rangle \tag{7.83}$$

with $b \in \mathbb{R}$ and $a^2 + b^2 = 1$. Resolution leads to

$$|P\rangle = a|r\rangle + b|g\rangle \tag{7.84}$$
$$|C\rangle = b|r\rangle - a|g\rangle\ . \tag{7.85}$$

If Alice or Bob is asked for the colour, he or she carries our a measurement on the subsystem in the $\{|P\rangle, |C\rangle\}$ basis and interprets the result with respect to the state to which the measurement leads, according to the rule $|P\rangle \leftrightarrow$ "peas!" and $|C\rangle \leftrightarrow$ "carrots!". Correspondingly, when the question is for the vegetable, the measurement is carried out in the rotated $\{|r\rangle, |g\rangle\}$ basis and the answer is given according to the rule $|r\rangle \leftrightarrow$ "red!" and $|g\rangle \leftrightarrow$ "green!". We can read off from the state $|\chi^{AB}\rangle$ the probabilities with which particular pairs of answers will be given.

To find the probabilities, we insert $|P\rangle$ from Eq. (7.84) in various places into Eq. (7.81):

$$|\chi^{AB}\rangle = N(|r^A,\, r^B\rangle - a^2(a|r^A\rangle + b|g^A\rangle)(a|r^B\rangle + b|g^B\rangle)) \tag{7.86}$$
$$|\chi^{AB}\rangle = N(|r^A,\, r^B\rangle - a^2(a|r^A\rangle + b|g^A\rangle)|P^B\rangle) \tag{7.87}$$
$$|\chi^{AB}\rangle = N(|r^A,\, r^B\rangle - a^2|P^A\rangle(a|r^B\rangle + b|g^B\rangle))\ . \tag{7.88}$$

$|r\rangle$ from Eq. (7.82) inserted into $|\chi^{AB}\rangle$ leads to:

$$\begin{aligned} |\chi^{AB}\rangle &= N[(a|P^A\rangle + b|C^A\rangle)(a|P^B\rangle + b|C^B\rangle) - a^2|P^A,\, P^B\rangle] \\ &= N(b^2|C^A,\, C^B\rangle + ab(|C^A,\, P^B\rangle + |P^A,\, C^B\rangle)) \,. \end{aligned} \tag{7.89}$$

From Eqs. (7.86)–(7.89), we find the probabilities for the possible pairs of measurement results and thus for the pairs of answers (compare Eq. (7.68)). From Eq. (7.86), for the pair of questions in the first case, the observed result is found:

$$p(g^A,\, g^B) = Na^4b^4 \neq 0 \,. \tag{7.90}$$

Eqs. (7.87) and (7.88) lead to

$$p(g^A,\, C^B) = 0, \quad p(C^A,\, g^B) = 0 \,. \tag{7.91}$$

Therefore, when the pair of questions (colour?, vegetable?) is asked, and Alice answers with "green!", then Bob always answers with "peas!" and *vice versa*. This reproduces the second case. Finally, we verify the third case with Eq. (7.89):

$$p(P^A,\, P^B) = 0 \,. \tag{7.92}$$

Entanglement is the tool with which quantum magicians can carry out the tricks which classical magicians cannot master.

Consequences The experimental implementations of the basic idea of the conjuring trick are not so simple as we described in Sect. 7.7.1. The results however agree well with the quantum theoretical predictions (compare Sect. 10.8).

The conjuring trick can be carried out in principle, but not with the means and methods of classical physics. *Systems for which the results of measurements (Alice's and Bob's answers) are already predetermined before the measurement (Einstein's reality) on the corresponding subsystems (Einstein's locality), such as is the case for the slips of paper, cannot be the cause of the observed correlations.* This shows that local-realistic theories and the quantum theory can lead to differing predictions. In our example: "the conjuring trick cannot be carried out" or "the conjuring trick can be carried out"; but only the predictions of quantum theory can be experimentally verified. Thus, local-realistic *alternative theories* to the quantum theory are refuted. We will describe additional experiments in Chap. 10 and then give more precise definitions for the concepts of *Einstein reality* and *Einstein locality*.

7.8 Quantum Gates for Multiple Qubit Systems

7.8.1 Entanglement via a CNOT Gate

The processing of quantum information is often explained schematically without reference to an experimental implementation with the aid of *quantum circuits*. The essential devices which are needed are: *quantum wires*; these are special quantum channels through which quantum systems can propagate without being modified; and *quantum gates*, which effect

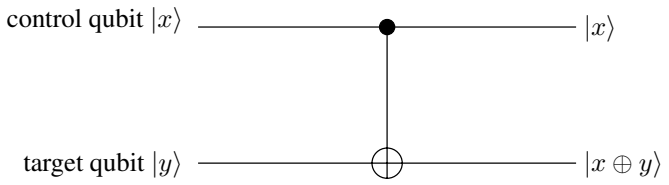


Figure 7.4: A CNOT gate.

unitary transformations of quantum systems. The systems are multi-qubits from the spaces $\mathcal{H}_2 \otimes \mathcal{H}_2 \otimes \mathcal{H}_2 \ldots \otimes \mathcal{H}_2$. *Measurements* permit reading out of the information. Owing to the unitarity of their operations, quantum gates represent reversible processes. Measurements are, in contrast, irreversible. *Quantum computers* are a network of quantum gates. We have already encountered quantum gates for quantum systems in $\mathcal{H}_2$ in Sect. 3.4. We now move on to product spaces. In Chap. 12, we will assemble quantum circuits into quantum computers.

Entanglement via a CNOT gate A simple quantum gate which transforms a qubit product state into an entangled state is the *CNOT gate* or controlled NOT gate, also called an XOR gate. Its action on the computational basis of $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$ is defined by

$$|x, y\rangle \to |x, y \oplus x\rangle \tag{7.93}$$

with $x, y, \ldots \in \{0, 1\}$. This determines the action on an arbitrary vector from $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$. The symbol $\oplus$ denotes addition modulo 2, i. e. $1 \oplus 1 = 0$. In detail, this means that:

$$|0^A, 0^B\rangle \overset{\text{CNOT}}{\longrightarrow} |0^A, 0^B\rangle \tag{7.94}$$

$$|0^A, 1^B\rangle \overset{\text{CNOT}}{\longrightarrow} |0^A, 1^B\rangle \tag{7.95}$$

$$|1^A, 0^B\rangle \overset{\text{CNOT}}{\longrightarrow} |1^A, 1^B\rangle \tag{7.96}$$

$$|1^A, 1^B\rangle \overset{\text{CNOT}}{\longrightarrow} |1^A, 0^B\rangle \; . \tag{7.97}$$

From this, it follows that

$$(\text{CNOT}) \cdot (\text{CNOT}) = \mathbb{1} \; . \tag{7.98}$$

Applying the matrix representation in the computational basis,

$$\text{CNOT} \leftrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \; , \tag{7.99}$$

one can readily verify the unitarity property:

$$(\text{CNOT})^\dagger = (\text{CNOT})^{-1} \; . \tag{7.100}$$

The qubits of the system A or B are called *control qubits* or *target qubits* (see Fig. 7.4).

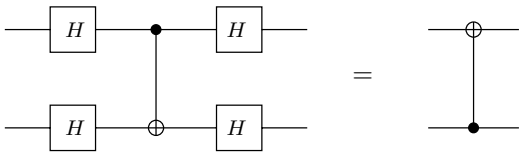


Figure 7.5: Two equivalent networks.

A simple example shows that *the CNOT gate transforms superpositions of control qubits into entanglements of control and target qubits:*

$$\left(\alpha|0^A\rangle \pm \beta|1^A\rangle\right)|0^B\rangle \quad \stackrel{\text{CNOT}}{\longrightarrow} \quad \alpha|0^A, 0^B\rangle \pm \beta|1^A, 1^B\rangle\,, \tag{7.101}$$

$$\left(\alpha|0^A\rangle \pm \beta|1^A\rangle\right)|1^B\rangle \quad \stackrel{\text{CNOT}}{\longrightarrow} \quad \alpha|0^A, 1^B\rangle \pm \beta|1^A, 0^B\rangle\,. \tag{7.102}$$

For $\alpha = \beta = \frac{1}{\sqrt{2}}$, in this manner four Bell states are formed. The reduced density operator of the target qubit is in this case $\rho^B = \frac{1}{2}\mathbb{1}^B$ (and correspondingly for the control qubit). Measurement in an arbitrary ONB of $\mathcal{H}_2^B$ yields the two measured values and states in perfect randomness with the probabilities $\frac{1}{2}$.

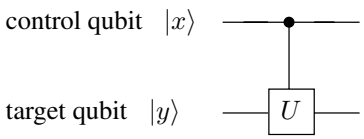


Figure 7.6: A controlled U gate.

A CNOT gate and four Hadamard gates can be combined to give the inverse of a CNOT gate (see Fig. 7.5). The CNOT gate is a special case of a *controlled U gate* (see Fig. 7.6). It leaves $|0, 0\rangle$ and $|0, 1\rangle$ unchanged. $|1, y\rangle$ with $y = 0, 1$ goes over to $|1\rangle \otimes U|y\rangle$. CNOT corresponds to $U = \sigma_x$.

7.8.2 Toffoli, SWAP, and Deutsch Gates

The *Toffoli gate* in Fig. 7.7 is also called a *CCNOT gate* (controlled-controlled NOT) or doubly-controlled NOT gate. In this case, the NOT gate acts on the target qubit if and only if both control qubits are in the state $|1\rangle$. The action of CCNOT is

$$|x, y, z\rangle \to |x, y, z \oplus xy\rangle\,. \tag{7.103}$$

The *SWAP gate* exchanges qubit states

$$\text{SWAP}|x^A, y^B\rangle = |y^A, x^B\rangle\,. \tag{7.104}$$

Analogously, one can construct a doubly-controlled U gate (see Fig. 7.8). It can be implemented with three CNOT gates (cf. Fig. 7.9):

$$|x, y\rangle \to |x, x \oplus y\rangle \to |y, x \oplus y\rangle \to |y, x\rangle\,. \tag{7.105}$$

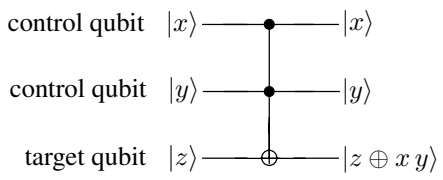


Figure 7.7: A Toffoli gate.

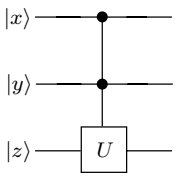


Figure 7.8: A doubly-controlled U gate.

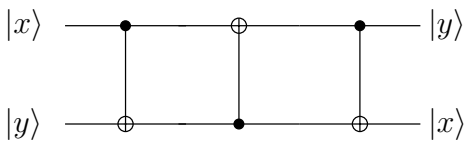


Figure 7.9: Exchange of two qubits (a SWAP gate).

Universal quantum gates are a series of quantum gates with which one can carry out every unitary transformation on $\mathcal{H}_2 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_2$. It can be shown that e. g. the *Deutsch gate* suffices for this purpose ([Deu 89]). In the case of this gate, the unitary transformation U in Fig. 7.8 has the form

$$U = -i \exp\left(i\frac{\theta}{2}\sigma_x\right) \ . \tag{7.106}$$

There are other universal gates (compare Sect. 7.10). We shall return at length to this topic in Sect. 12.9.

7.9 Systems of Identical Particles*

In connection with systems whose subsystems contain elementary particles of the same type – we consider as an example two spin-$\frac{1}{2}$ particles – the following questions are frequently asked:

*The sections marked with an asterisk * can be skipped over in a first reading.

The particles are Fermions and their composite states must be antisymmetric with respect to exchange of the states of the individual particles. Therefore, they must take a form similar to the Bell vectors $|\Phi_-\rangle$ or $|\Psi_-\rangle$. Why can we not construct e.g. a teleportation procedure based on this always-present "natural" entanglement? And conversely, how can we implement teleportation with Fermions in a symmetrically-entangled state $|\Phi_+\rangle$?

Identical particles *Identical particles* have the same values of all their *intrinsic properties* such as mass, charge, spin etc. Electrons for example can be distinguished from positrons but not from each other. They cannot be marked and therefore have no individuality. An identification is not possible.

To describe systems of identical particles, one can begin with enumerated distinguishable particles and then remove their distinguishability. We restrict our considerations to 2-particle systems. The generalisation to more particles is straightforward. The state vectors of two distinguishable particles with the numbers (1) and (2) lie in the product space $\mathcal{H}^{(1)(2)} = \mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)}$. According to the postulates, the states of identical particles are either completely symmetric in their particle numbers (Bosons, with integral spins), or completely antisymmetric (Fermions, with half-integral spins). Their state vectors lie correspondingly in subspaces of $\mathcal{H}^{(1)(2)}$, which we denote by $\mathcal{H}_+^{(1)(2)}$ and $\mathcal{H}_-^{(1)(2)}$. These subspaces are themselves not product spaces. If $\{|u^{(1)}\rangle\}$ and $\{|v^{(2)}\rangle\}$ are ONB of $\mathcal{H}^{(1)}$ or $\mathcal{H}^{(2)}$, respectively, then the bases of e.g. $\mathcal{H}_-^{(1)(2)}$ are given by $|n,i\rangle_- := \frac{1}{\sqrt{2}}(|n^{(1)},i^{(2)}\rangle - |i^{(1)},n^{(2)}\rangle)$ with $1 \leq n \leq \dim\mathcal{H}^{(1)}$ and $1 \leq i \leq \dim\mathcal{H}^{(2)}$. Without any interactions at all, the symmetry postulate leads to states which are formally entangled in terms of their non-observable particle numbers.

Even with the aid of observables, no identification of the particles is allowed. Observables must therefore be invariant under permutations of the particle numbers. The postulates for projection measurements apply. They are formulated with respect to the spaces $\mathcal{H}_+^{(1)(2)}$ or $\mathcal{H}_-^{(1)(2)}$. A consequence of this is that measurements or unitary dynamical evolutions cannot produce transitions between Bosons and Fermions.

Particles in two different regions of space We clarify the essential points using the example of two spin-$\frac{1}{2}$ particles with external degrees of freedom. $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$ are thus already assumed to be product spaces for the external degrees of freedom (vectors $|\alpha\rangle$ and $|\beta\rangle$) and for the spins (vectors $|0\rangle$ and $|1\rangle$). A possible state is then e.g.

$$|\Lambda^{(1)(2)}\rangle = \frac{1}{\sqrt{2}}\left(|\alpha,0\rangle^{(1)} \otimes |\beta,1\rangle^{(2)} - |\beta,1\rangle^{(1)} \otimes |\alpha,0\rangle^{(2)}\right) . \tag{7.107}$$

In order to make the connection to the preceding sections, we discuss a situation in which $\langle\alpha|\beta\rangle = 0$ holds. The states $|\alpha\rangle$ and $|\beta\rangle$ are orthogonal. This is the case e.g. when the particles have differing directions of their momenta or when the wavefunctions $\langle\vec{r}|\alpha\rangle$ and $\langle\vec{r}|\beta\rangle$ are nonzero only in a restricted spatial region G_α or G_β, where G_α and G_β do not overlap ($G_\alpha \cap G_\beta = \emptyset$). Then, one can register a Fermion only in G_α or in G_β, but not outside them. However, statements about the particle numbers (1) or (2) are not possible. G_α or G_β can be e.g. different locations at which Alice A or Bob B have set up their measurement apparatus

which can carry out measurements in spin space. If Alice's apparatus registers a signal, this at the same time represents a measurement in configuration space, i.e. G_α is registered. For the description of this situation, we can introduce an abbreviated form which reflects the fact that in the case $G_\alpha \cap G_\beta = \emptyset$, the state $|\alpha\rangle$ (i.e. the location G_α, Alice) is always correlated with $|0\rangle$ and the state $|\beta\rangle$ (i.e. the location G_β, Bob) is always correlated with $|1\rangle$:

$$|\Lambda^{(1)(2)}\rangle \leftrightarrow |\Lambda^{AB}\rangle = |0^A, 1^B\rangle \,. \tag{7.108}$$

With respect to all the measurements which can be carried out by Alice and Bob, the product state $|\Lambda^{AB}\rangle$ is equivalent to the state $|\Lambda^{(1)(2)}\rangle$. If Alice measures the observable σ_z, she always finds the spin state $|0\rangle$. This is the content of Eq. (7.107).

If Alice measures the observable σ_x, the result of the measurement can yield e.g. the eigenvalue $|1_x\rangle$ and the 2-Fermion system is transformed after selection into the state

$$|\Lambda^{AB}\rangle \to |\Lambda'^{AB}\rangle = |1_x^A, 1^B\rangle \,. \tag{7.109}$$

$|\Lambda'^{AB}\rangle$ can again be written in the complete form $|\Lambda'^{(1)(2)}\rangle$. To this end, we replace $|0\rangle$ on the right-hand side of Eq. (7.107) by $|1_x\rangle$. The probability of this result is $|\langle\Lambda^{(1)(2)}|\Lambda'^{(1)(2)}\rangle|^2 = |\langle\Lambda^{AB}|\Lambda'^{AB}\rangle|^2$. As a result of the orthonormalisation $\langle\alpha|\beta\rangle = 0$, the vector $|\Lambda^{AB}\rangle$ in Eq. (7.108) is a product vector in $\mathcal{H}^{AB}$.

Utilisable and non-utilisable entanglement The state introduced above, $|n, i\rangle_-$, is a superposition, from which measurable interference effects can result in particular physical situations. The fact that the particles cannot be distinguished has physical consequences. The energy spectrum of the helium atom is an example of this. In the following sections, however, we shall discuss other physical questions. The entanglement for example in the state $\frac{1}{\sqrt{2}}(|0^{(1)}, 1^{(2)}\rangle - |1^{(1)}, 0^{(2)}\rangle)$ is related to the indistinguishable particle numbers. They do not denote subsystems. Since this formal entanglement cannot be used, it cannot serve as a tool for quantum-mechanical information processing. It is not utilisable for this purpose.

Only the entanglement with the states $|\alpha\rangle$ and $|\beta\rangle$ with $\langle\alpha|\beta\rangle = 0$, as in the state $|\Lambda^{(1)(2)}\rangle$ of Eq. (7.107), opens up the possibility of intercession via $|\alpha\rangle$ and $|\beta\rangle$. As we have already seen, then $|\Lambda^{(1)(2)}\rangle$ becomes equivalent to a non-entangled product state $|\Lambda^{AB}\rangle$. If we now form e.g. by superposition with an additional state

$$|\Omega^{(1)(2)}\rangle = \frac{1}{\sqrt{2}}\left(|\alpha, 1\rangle^{(1)} \otimes |\beta, 0\rangle^2 - |\beta, 0\rangle^{(1)} \otimes |\alpha, 1\rangle^{(2)}\right) \leftrightarrow |\Omega^{AB}\rangle = |1^A, 0^B\rangle \tag{7.110}$$

the state vector

$$|\Psi_+^{(1)(2)}\rangle := \frac{1}{\sqrt{2}}\left(|\Lambda^{(1)(2)}\rangle + |\Omega^{(1)(2)}\rangle\right) \,, \tag{7.111}$$

then we can read off from the abbreviated notation

$$|\Psi_+^{(1)(2)}\rangle \leftrightarrow |\Psi_+^{AB}\rangle = \frac{1}{\sqrt{2}}\left(|0^A, 1^B\rangle + |1^A, 0^B\rangle\right) \tag{7.112}$$

that the Bell state $|\Psi_+^{AB}\rangle$ has been formed. In spite of the addition in Eq. (7.112), it has the symmetry properties required for the description of two identical Fermions. At the same time, a utilisable entanglement has come about by the superposition described by Eq. (7.111).

If the condition $\langle\alpha|\beta\rangle = 0$ is not fulfilled in a physical situation, it can be expected that additional effects will occur in the course of the information processing, which are due to the indistinguishability of the particles. One can also switch on and off the coupling $\langle\alpha|\beta\rangle \neq 0$ in the form of a time-dependent *exchange coupling* and thereby produce an entanglement only at certain times. We mention also that the symmetry or antisymmetry property is automatically taken into account within the framework of the *second quantisation*.

7.10 Complementary Topics and Further Reading

- For "proper mixtures" and "improper mixtures": [d'Es 95], [d'Es 99].
- Local measurements and the requirements of the theory of relativity: [PT 04].
- The idea that the whole is more than the sum of its parts is referred to in philosophy as *holism*. There is a whole series of philosophical analyses in which the attempt is made to give this idea a precise meaning in many different fields from sociology to physics, and to investigate its consequences. For the natural-philosophical question as to whether there is a holism in physics, quite new aspects have resulted from the study of entangled states in composite systems (see [Pri 81, Sects. 3.7, 5.6, 6.3]). Two differing analyses of this question are introduced in [Esf 04] and [See 04] (cf. [Esf 06]). There, more detailed literature is also cited. See also [Hea 99].
- For the conjuring trick: [Har 93], [Har 98].
- Experiments on the conjuring trick: [Har 92], [TBM 95], [DMB 97], [BBD 97].
- An overview of quantum gates for qubits: [DiV 98], [Bra 02].
- In utilising coupled *quantum dots* or neutral atoms in *microtraps* as tools for quantum information processing, effects occur which are based upon the indistinguishability of particles. For details and further literature, see: [ESB 02].

7.11 Problems for Chapter 7

Prob. 7.1 [for 7.3.2]: Show that ρ^A and ρ^B in Eq. (7.34) have the properties required of a density operator.

Prob. 7.2 [for 7.4 and 7.5]: Confirm the results of Sects. 7.4 and 7.5 for the case that the initial state was not a pure state $|\psi^{AB}\rangle$, but rather a mixture, ρ^{AB}.

Prob. 7.3 [for 7.5.2]: Prove Eq. (7.75).

Prob. 7.4 [for 7.8]: Show in each case the equivalence of the networks asserted in Figs. 7.5 and 7.10.

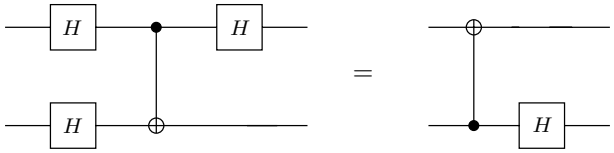


Figure 7.10: Two equivalent networks.

Prob. 7.5 [for 7.8:] Show that the network in Fig. 7.11 converts pairs of Bell states into pairs of Bell states.

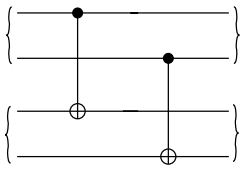


Figure 7.11: Mapping of Bell states onto Bell states.

Prob. 7.6 [for 7.8]: Show that one can construct a *quantum adder* from a Toffoli gate and a CNOT gate.

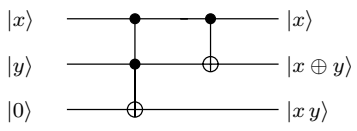


Figure 7.12: A quantum adder. The first two bits are added modulo 2. The circuitry is reversible.

© 2007 WILEY-VCH Verlag GmbH & Co.

8 Entanglement

In this chapter, we want to establish operationally the concept of quantum correlations and distinguish them from classical correlations, which can also be present in composite quantum systems. We also will establish the connection between quantum correlations and entanglement. For pure states, the presence of entanglement can be determined by using the Schmidt decomposition. A measure of the entanglement will be introduced, and an example of the creation of entangled states will be given. For applications, it is important that quantum states cannot be copied. The fact that states can be marked through their entanglement with other states leads to the quantum eraser and to the question of "delayed choice".

8.1 Correlations and Entanglement

A composite spin system with two subsystems can be e.g. in the product states $|0^A, 0^B\rangle$ or $|1^A, 1^B\rangle$, or also in their superpositions $\alpha|0^A, 0^B\rangle + \beta|1^A, 1^B\rangle$. The superposition $\alpha \neq 0, \beta \neq 0$ is an example of an entangled state. Entangled states play a fundamental role in quantum information processing. They are the central tool with which non-classical effects can be produced.

Composite systems in entangled states are correlated. If, for the state given above, the observable σ_z is measured on each of the subsystems, then one finds as the combination of measured values either $(-1, -1)$ or $(+1, +1)$. In comparison to correlated classical systems, the correlations have in this case a different structure. That can be seen if one includes measurements of other observables besides σ_z. We shall discuss this point in detail in Chap. 10. In the following sections, we will explain the relevant concepts more specifically, and we start directly with mixtures.

8.1.1 Classically-Correlated Quantum States and LOCC

Correlated quantum states We begin with the fundamental distinction between correlated and non-correlated quantum states (see Fig. 8.1). We consider again a composite system S^{AB} with the subsystems S^A and S^B. As shown by the considerations in Chap. 7, in a product state $\rho^{AB} = \rho^A \otimes \rho^B$, the subsystems are in every respect completely independent of each other. They are not correlated. Conversely, we shall call states ρ^{AB} of composite systems which are not product states

$$\rho^{AB} \neq \rho^A \otimes \rho^B \,, \tag{8.1}$$

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

correlated. Note that the use of the same core symbol ρ is not intended to imply that the states of S^A and S^B are the same.

This mathematical characterisation is equivalent to an operational statement about measured values which in principle can be verified by measurements on subsystems (cf. Sect. 7.5.1): a state ρ^{AB} is correlated if and only if there are observables C^A and D^B for which the expectation value of $C^A \otimes D^B$ is not given by the product of the expectation values of the reduced density operators (compare problem 8.1 in Sect. 8.9):

$$\mathrm{tr}[(C^A \otimes D^B)\rho^{AB}] \neq \mathrm{tr}_A[C^A\rho^A] \cdot \mathrm{tr}_B[D^B\rho^B]\,. \tag{8.2}$$

This agrees with the intuitive interpretation.

Classically-correlated quantum states Quantum systems can be in states which lead to correlations in the measured values which are similar to the correlations observed for classical systems. We want to first give a more detailed description of these classically-correlated quantum states. In a second step, we characterise the states whose correlations lead to the non-classical effects which are important for our topic. We however will not start in this case with the measured values, but rather with the preparation procedure.

The composite system S^{AB} is supposed to be prepared by Alice or Bob *starting from product states* by means of *local operations* (LO) on the subsystems S^A and S^B. Among such operations are unitary dynamic actions, measurements, and all other local manipulations[1]. Furthermore, Alice and Bob can exchange information via *classical communication* (CC) such as by telephone etc., in order to coordinate their local operations. We abbreviate *"local operations and classical communication"* as *LOCC*.

In this preparation of the composite system by means of LOCC, Alice prepares the system S^A in the state ρ_r^A and informs Bob of this via classical communication channels; he, for his part, prepares his system S^B in the state ρ_r^B which is in general different from that prepared by Alice. This sequence of actions is repeated many times for different states in a random fashion, keeping track of relative frequencies p_r. The composite state thus prepared, ρ^{AB}, is then by construction a convex combination or a statistical mixture (blend) of product states

$$\rho^{AB} = \sum_{r=1}^{m} p_r \rho_r^A \otimes \rho_r^B \quad , \quad p_r > 0 \quad , \quad \sum_r p_r = 1\,. \tag{8.3}$$

Such a procedure could be carried out in a similar manner with classical states. Comparison with Eq. (8.1) shows that the resulting state ρ^{AB} is correlated, if the sum cannot be reduced to a single term. Since the correlations are produced by LOCC in a purely classical manner with the probabilities p_r, one terms a quantum state ρ^{AB}, which can be written as a genuine mixture ($m \neq 1$) of product states, a *classically-correlated state*. We also note that after the introduction of ensemble decompositions for all the ρ_r^A and ρ_r^B, the density operator can also be written in the form

$$\rho^{AB} = \sum_j \pi_j |a_j^A\rangle\langle a_j^A| \otimes |b_j^B\rangle\langle b_j^B|\,, \tag{8.4}$$

with $0 < \pi_j \leq 1$ and $\sum_j \pi_j = 1$. The states involved need not be orthogonal.

[1] These can also be generalised measurements, for which additional ancillary systems are used (cf. Chap. 13).

As in the transition between mixtures, we again disengage ourselves from the particular preparation procedure. A quantum state ρ^{AB} is termed classically correlated even if it was not produced by the preparation procedure described above. It suffices that it can be simulated in every respect by a state prepared in this manner, i.e. that its statistical properties can be reproduced by a LOCC mechanism. Mathematically, this means that ρ^{AB} can be written in the form (8.4).

The term "classically" correlated with respect to quantum states should not give rise to misconceptions. The analogy to correlated classical states is not perfect. We again consider the example of a set of many pairs of boxes which both contain either red or blue balls with the probabilities p_1 or p_2, respectively ($p_1 + p_2 = 1$). When the boxes are opened, one can observe a correlation of the colours, which is analogous to the correlation in the measurement of the spin components in the z-direction in a classically-correlated quantum state

$$\begin{aligned} \rho^{AB} &= p_1|0^A, 0^B\rangle\langle 0^A, 0^B| + p_2|1^A, 1^B\rangle\langle 1^A, 1^B| \\ &= p_1|0^A\rangle\langle 0^A| \otimes |0^B\rangle\langle 0^B| + p_2|1^A\rangle\langle 1^A| \otimes |1^B\rangle\langle 1^B| \,. \end{aligned} \tag{8.5}$$

The states $|0\rangle$ and $|1\rangle$ correspond to the two colours. If one carries out a measurement in the computational basis on the subsystem S^A, then it leads with the probability p_1 to the state $|0^A\rangle$, and S^B is found after the selective measurement in the state $|0^B\rangle$ which is correlated with $|0^A\rangle$.

However, one can also carry out the measurement in a basis which is "rotated" with respect to the basis $\{|0\rangle, |1\rangle\}$. There is no analogy for the coloured balls. A measurement on the system in the state ρ^{AB} of Eq. (8.5) with $p_1 = p_2 = \frac{1}{2}$ in the basis $\{|0^A_x\rangle, |1^A_x\rangle\}$, which leads to the final state $|0^A_x\rangle$ for S^A, gives rise for the composite state to

$$\rho^{AB} \to \rho^{AB'} = |0^A_x\rangle\langle 0^A_x| \otimes \frac{1}{2}\mathbb{1}^B \,. \tag{8.6}$$

The state of S^B is a maximally mixed state and in this sense it is completely undetermined. The results of the measurements of σ_x on both subsystems are uncorrelated.

8.1.2 Separability and Entanglement

It has proved to be helpful to introduce the following concepts (see Fig. 8.1): A state ρ^{AB} of a composite systems S^{AB} is called *separable* when it can be written in the form (8.3) of a convex combination of product states. A separable state is thus classically correlated or even non-correlated ($m = 1$). A pure or mixed quantum state which is *not* separable is called *entangled. An entangled quantum state thus contains non-classical correlations, which are also called* quantum correlations *or EPR correlations*[2].

This is the reason for their considerable physical significance. The preparation procedure LOCC leads to separable states. *Entangled states can not be produced from product states via*

[2]The acronym EPR is an abbreviation for the names of A. Einstein, B. Podolsky and N. Rosen. They were the authors of an important paper [EPR 35] which triggered the considerations that we describe in particular in Chap. 10. The term EPR has in the meantime become to a large extent independent of this original article and is used to denote for example the correlations which can be observed in entangled systems, and the corresponding experiments. EPR is thus today a part of the systematic terminology and not an historical reference.

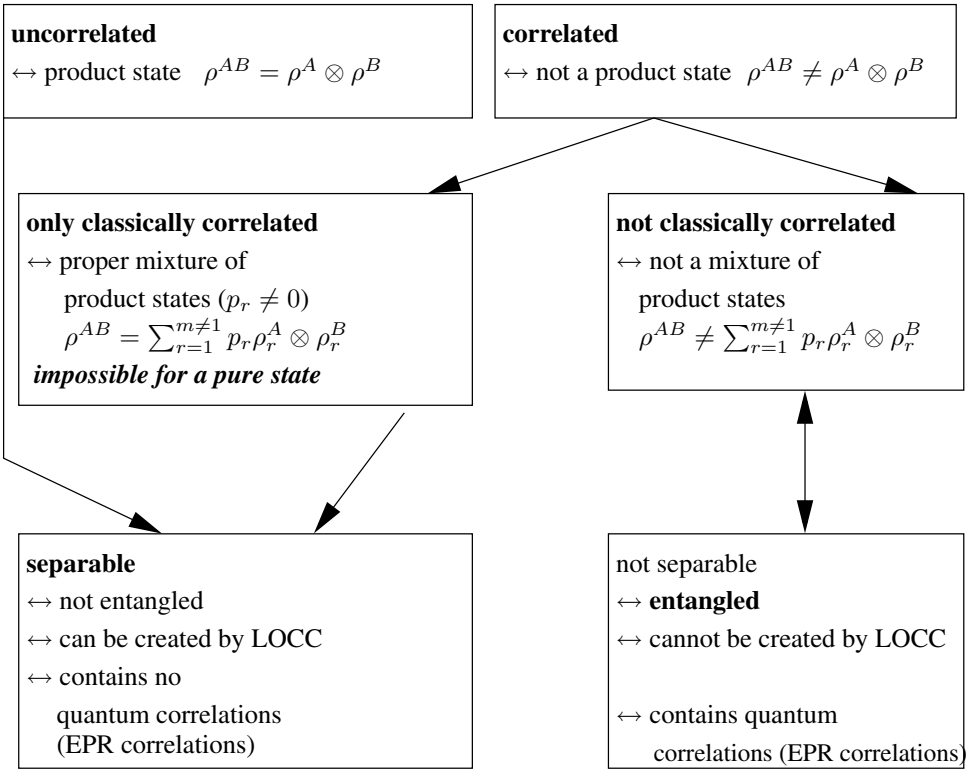

uncorrelated
↔ product state $\rho^{AB} = \rho^A \otimes \rho^B$
correlated
↔ not a product state $\rho^{AB} \neq \rho^A \otimes \rho^B$
only classically correlated
↔ proper mixture of
product states ($p_r \neq 0$)
$\rho^{AB} = \sum_{r=1}^{m\neq 1} p_r \rho_r^A \otimes \rho_r^B$
impossible for a pure state
not classically correlated
↔ not a mixture of
product states
$\rho^{AB} \neq \sum_{r=1}^{m\neq 1} p_r \rho_r^A \otimes \rho_r^B$
separable
↔ not entangled
↔ can be created by LOCC
↔ contains no
quantum correlations
(EPR correlations)
not separable
↔ entangled
↔ cannot be created by LOCC
↔ contains quantum
correlations (EPR correlations)

Figure 8.1: Distinction of cases for the characterisation of the states of bipartite systems. The converging arrows lead to the umbrella term. Diverging arrows point to special cases. The double arrow indicates an equivalence. A pure overall state is either not correlated or it is entangled.

LOCC. This characterisation can also be considered to be an equivalent definition of entanglement. The fact that non-local effects can be obtained via LOCC for already entangled states (e. g. enhancement of the entanglement) will be discussed in Sect. 13.3.6.

It is an important property of quantum physics as compared to classical physics that not all correlations are necessarily classical. This can be seen immediately from the example of the pure states in $\mathcal{H}^A \otimes \mathcal{H}^B$. We learned in Sect. 4.1.3 that the density operator for a pure state $|\psi^{AB}\rangle$ cannot be decomposed in terms of a convex sum. Thus, it cannot be classically correlated as in Eq. (8.3). *A pure state is either not correlated (then it is a product state), or it is EPR-correlated and thus entangled.* It then contains a type of correlations which does not occur in classical systems. We will clarify this point again from a different viewpoint in Sect. 9.2.4.

8.1.3 The Separability Problem

Since quantum correlations play a significant role as tools, it is an important goal to develop criteria with which we can read off whether and to what extent a given state is entangled. For pure states of bipartite systems, this is simple. *The state $|\psi^{AB}\rangle$ is separable if and only if it has the form $|\psi^{AB}\rangle = |\phi^A\rangle \otimes |\chi^B\rangle$, i.e. if each of its two reduced density operators is a pure state.* In Sect. 8.3, we will encounter a further criterion for pure states in terms of the Schmidt decomposition, which at the same time leads to a measure of the entanglement. Making use of the quantum entropy of the subsystems, we introduce in Sect. 8.3.3 the measure of the entanglement of pure states of bipartite systems which is at present generally accepted.

Realistic states, as are used for experiments, are nearly always mixed. They contain perturbing "admixtures" to the pure states which would in fact be needed. Furthermore, quantum systems cannot be perfectly decoupled from their environment. Therefore, additional subsystems take part in the composite system. Thus, a very complex situation arises, which can be only approximately described in the way we do in the following chapters. In addition, it is not very simple to arrive at an intuitive physical understanding of the entanglement of mixtures. We will elucidate this with an example. The Bell states of Eq. (7.12) are entangled pure states. Using the formal criterion given above, this can be seen directly from the reduced density operators, which are given by $\frac{1}{2}\mathbb{1}$. By mixing the entangled states $|\Phi_+^{AB}\rangle$ and $|\Phi_-^{AB}\rangle$ with equal probabilities, $\frac{1}{2}$, we obtain the state

$$\rho^{AB} = \frac{1}{2}(|\Phi_+^{AB}\rangle\langle\Phi_+^{AB}| + |\Phi_-^{AB}\rangle\langle\Phi_-^{AB}|)\,. \tag{8.7}$$

If we insert the definitions of the Bell states, we find

$$\begin{aligned}\rho^{AB} &= \frac{1}{2}(|0^A,0^B\rangle\langle 0^A,0^B| + |1^A,1^B\rangle\langle 1^A,1^B|) \\ &= \frac{1}{2}\left(|0^A\rangle\langle 0^A| \otimes |0^B\rangle\langle 0^B| + |1^A\rangle\langle 1^A| \otimes |1^B\rangle\langle 1^B|\right)\,. \end{aligned} \tag{8.8}$$

The state ρ^{AB} is of the type (8.3) and therefore purely classically correlated. It is not entangled. It is possible to prepare it alternatively by LOCC alone. In the original preparation procedure of ρ^{AB} according to Eq. (8.7), either the state $|\Phi_+^{AB}\rangle$ or the state $|\Phi_-^{AB}\rangle$ is produced. It is plausible that simply because of the random alternation of the different entangled states, their entanglement cannot for example be used to carry out the conjuring trick from Sect. 7.7. Presumably, no longer are there sufficient quantum correlations remaining after such a mixing process. We have shown that in fact, none at all remain.

The density operator of a given mixture has infinitely many ensemble decompositions. If for a bipartite system at least one of these has the form (8.3), then the state is separable; otherwise, it is entangled. For systems with many subsystems and Hilbert spaces of large dimensions, separability must make a statement about all the ensemble decompositions. This has not been satisfactorily formulated up to now. The relevant research programme is called the *separability problem*: is a state of a composite quantum system given by a generalised density operator separable or not, and how can one determine this experimentally?

In simple cases, however, it is possible to give a measure of the degree of correlation of two systems. We will describe the degree of correlation of pure states in Chap. 9 with the

help of von Neumann's mutual entropy (or information). This is a wide field in which the entropy plays an important role. We discuss the PPT criterion in Sect. 8.4, and indicators of entanglement in Sect. 10.6. In Section 11.7, we encounter the concurrence as a measure of the entanglement of two qubit systems.

8.2 Outlook

In Sect. 2.2, we gave an overview which on the main referred to the two types of dynamics in quantum physics: the deterministic dynamics between preparation and measurement, and the non-deterministic dynamics of the measurement process. In the chapters that followed, we encountered the von Neumann entropy and the concept of entanglement for mixtures. A setup for the production of entanglement was described, and – with the conjuring trick – an example was given of its application. In the coming chapters, we want to complement the overviews in Sects. 2.2 and 7.3.1 by treating *entanglement* as a focal-point subject. This leads us to the following topics:

(i) ***Understanding*** *entanglement*

- The marking of states
- Entropy, types of correlations, non-local measurements
- Exchange of entanglement
- Decoherence

(ii) ***Producing*** *entanglement*

- Cascade photons

(iii) ***Using*** *entanglement*

- Refuting local-realistic theories
- Quantum cryptography
- Dense coding
- Quantum teleportation
- Quantum computers
- Implementation of quantum operations and generalised measurements

(iv) ***Detecting, quantifying****, and* ***concentrating*** *entanglement*

- Separability criteria (e.g. the PPT criterion)
- Indicators of entanglement for experimental detection
- Concurrence as a measure of entanglement
- Distilling entanglement

We begin with a discussion of the situation in the case of pure states.

8.3 Entangled Pure States

8.3.1 The Schmidt Decomposition

For a discussion of the entanglement of pure states $|\psi^{AB}\rangle$ of bipartite systems, it has proved especially helpful to make use of the *Schmidt decomposition*. It is also called the *bi-orthogonal or polar decomposition* (bi-orthogonal or polar expansion) of the vector $|\psi^{AB}\rangle$. It states the following:

Let $|\psi^{AB}\rangle$ *be a normalised pure state of the composite system* S^{AB} *in the product Hilbert space* $\mathcal{H}^{AB} = \mathcal{H}^A \otimes \mathcal{H}^B$ *with* $\dim \mathcal{H}^A = a$ *and* $\dim \mathcal{H}^B = b$. *With* $\rho^{AB} = |\psi^{AB}\rangle\langle\psi^{AB}|$, *the operators* $\rho^A = tr_B[\rho^{AB}]$ *and* $\rho^B = tr_A[\rho^{AB}]$ *are the reduced density operators of the subsystems* S^A *and* S^B. *Then we have the following results:*

(i) *The vector* $|\psi^{AB}\rangle$ *can be written in the form of a* Schmidt decomposition[3]

$$|\psi^{AB}\rangle = \sum_{n=1}^{k} \sqrt{p_n}|u_n^A, w_n^B\rangle \quad \text{with} \quad p_n > 0 \tag{8.9}$$

with $k \leq \min(a, b)$, *where* $\{|u_n^A\rangle\}$ *and* $\{|w_n^B\rangle\}$ *are the orthonormalised eigenvectors of* ρ^A *in* $\mathcal{H}^A$ *(or* ρ^B *in* $\mathcal{H}^B$*) with suitably chosen phases. For pairwise differing eigenvalues* p_n *(i.e. no degeneracy), the vectors* $|u_n^A\rangle$ *and* $|w_n^B\rangle$ *are uniquely determined up to a phase. It follows from this that:*

(ii) ρ^A *and* ρ^B *have the same positive eigenvalues* p_1,...,p_k *(for g-fold degeneracy, the corresponding eigenvalue is to be repeated g times).*

The number k is called the *Schmidt rank* of $|\psi^{AB}\rangle$.

To prove (i), we expand $|\psi^{AB}\rangle$ in the ONB $\{|n^A\rangle, 1 \leq n \leq a\}$ of $\mathcal{H}^A$ and $\{|i^B\rangle, 1 \leq i \leq b\}$ of $\mathcal{H}^B$ with $a \leq b$

$$|\psi^{AB}\rangle = \sum_{n,i=1}^{a,b} a_{ni}|n^A, i^B\rangle \tag{8.10}$$

and again introduce the relative states

$$|\tilde{w}_n^B\rangle := \sum_{i=1}^{b} a_{ni}|i^B\rangle \tag{8.11}$$

leading to

$$|\psi^{AB}\rangle = \sum_{n=1}^{a} |n^A, \tilde{w}_n^B\rangle\,. \tag{8.12}$$

[3] E. Schmidt, Math. Ann. 63, 433-476 (1907).

The relative states $|\tilde{w}_n^B\rangle$ are in general neither orthogonal nor normalised. We show how the orthogonality claimed for Eq. (8.9) can be obtained.

To do this, we choose as the ONB $\{|n^A\rangle\}$ in Eq. (8.10) the particular orthonormal eigenvectors $\{|u_n^A\rangle\}$ of ρ^A and expand ρ^A

$$\rho^A = \sum_{n=1}^{a} p_n |u_n^A\rangle\langle u_n^A| \quad \text{with} \quad p_n \geq 0\,, \quad \sum_{n=1}^{a} p_n = 1\,. \tag{8.13}$$

Let $p_n > 0$ for $1 \leq n \leq k$ and $p_n = 0$ for $k+1 \leq n \leq a$. The eigenvalues p_n can be degenerate. The vector $|u_n^A\rangle$ is determined only up to a phase for non-degenerate eigenvalues. On the other hand, with the relative states associated with $|u_n^A\rangle$, $\{|\tilde{w}_n^B\rangle\}$, which we shall use from now on, we find

$$\begin{aligned}
\rho^A &= \mathrm{tr}_B[|\psi^{AB}\rangle\langle\psi^{AB}|] = \mathrm{tr}_B\left[\sum_{l,n}^{a} |u_l^A, \tilde{w}_l^B\rangle\langle u_n^A, \tilde{w}_n^B|\right] = \\
&= \mathrm{tr}_B[\sum_{l,n=1}^{a} |u_l^A\rangle\langle u_n^A| \otimes |\tilde{w}_l^B\rangle\langle\tilde{w}_n^B|] = \\
&= \sum_{l,n=1}^{a} |u_l^A\rangle\langle u_n^A| \mathrm{tr}_B[|\tilde{w}_l^B\rangle\langle\tilde{w}_n^B|] = \\
&= \sum_{l,n=1}^{a} \langle\tilde{w}_n^B|\tilde{w}_l^B\rangle |u_l^A\rangle\langle u_n^A|\,,
\end{aligned} \tag{8.14}$$

where in the last step (compare Eq. (1.35) with $A = \mathbb{1}$)

$$\mathrm{tr}_B[|\tilde{w}_l^B\rangle\langle\tilde{w}_n^B|] = \sum_{i=1}^{b} \langle i^B|\tilde{w}_l^B\rangle\langle\tilde{w}_n^B|i^B\rangle \tag{8.15}$$

$$= \sum_{i=1}^{b} \langle\tilde{w}_n^B|i^B\rangle\langle i^B|\tilde{w}_l^B\rangle = \langle\tilde{w}_n^B|\tilde{w}_l^B\rangle \tag{8.16}$$

was inserted. Comparison of (8.13) with (8.14) leads to the orthogonality of the relative states:

$$\langle\tilde{w}_n^B|\tilde{w}_l^B\rangle = p_n \delta_{nl}\,. \tag{8.17}$$

For $n \geq k+1$, the $|\tilde{w}_n^B\rangle$ are null vectors. If one furthermore takes into account the fact that we have chosen $\{|u_n^A\rangle\}$ as ONB instead of $\{|n^A\rangle\}$, we find as a result of Eq. (8.12) with (8.17) the proposition (i). Claim (ii) is a direct result.

From Eq. (8.9), the reduced density operator of S^A is found:

$$\rho^A = \sum_{i=1}^{k} p_n |u_n^A\rangle\langle u_n^A|\,. \tag{8.18}$$

The p_n are called *Schmidt coefficients*. We add that also the reduced density operator ρ^B has the same p_n as eigenvalues

$$\rho^B = \sum_{n=1}^{k} p_n |w_n^B\rangle\langle w_n^B|\,. \tag{8.19}$$

This has the immediate consequence that every function of a density operator which depends only on its eigenvalues has the same value for both reduced density operators. The degree of mixtureness from Sect. 4.1.3 and the von Neumann entropy, which we introduced in Chap. 6, are examples of this. They are the same for both subsystems.

A Schmidt decomposition always refers to a particular pure state of a composite system. Different states have differing Schmidt decompositions. In general, the Schmidt decomposition cannot be extended to systems with more than two subsystems. For mixtures, there is no analogue of the Schmidt decomposition.

8.3.2 The Schmidt Number and Entanglement

We will give some examples for the usefulness of the Schmidt decomposition. $\{|u_n^A\rangle\}$ or $\{|w_n^B\rangle\}$ with $n = 1, ..., k$ are called the *Schmidt bases* of $\mathcal{H}^A$ or $\mathcal{H}^B$. In them, the two reduced density operators are diagonal. The Schmidt decomposition is a superposition of separable states. All the information about the entanglement of $|\psi^{AB}\rangle$ is contained in the Schmidt coefficients. The *Schmidt number* k is the number of non-vanishing Schmidt coefficients. *$|\psi^{AB}\rangle$ is a product state and therefore not entangled, if and only if the Schmidt number is equal to one. ρ^A and ρ^B are then said to have the rank one. This is equivalent to* $tr[(\rho^A)^2] = tr[(\rho^B)^2] = 1$. Whether a pure state $|\psi^{AB}\rangle$ is entangled or not can thus be directly read off the reduced density operator of a subsystem. The Schmidt number can serve as a *measure of the entanglement of pure states*. This is the reason for its significance. Furthermore, we can read off directly: if one of the reduced density operators describes a pure state, then so does the other. *When the overall system S^{AB} is in a pure state $|\psi^{AB}\rangle$, it is thus impossible that one of the subsystems be in a pure state and the other in a genuine mixture.* We will generalise this statement further.

Similarly, we find that: when a qubit system is entangled with a system with m linearly-independent states, then the Schmidt decomposition contains only two terms. For calculations, it is often expedient to introduce the Schmidt basis. *In general, the statement holds: If a subsystem has the dimension d, then it cannot be entangled with more than d orthogonal states of another system.*

If ρ^A and thus also ρ^B have as their only degenerate eigenvalue at most zero, then the Schmidt decomposition is uniquely determined by ρ^A and ρ^B. One can find the corresponding eigenstates of ρ^A and ρ^B and form the product of the states belonging to the same eigenvalue as in Eq. (8.9). For $\dim \mathcal{H}^A \leq \dim \mathcal{H}^B$, the spectrum of ρ^B consists of the Schmidt coefficients and $(\dim \mathcal{H}^B - \dim \mathcal{H}^A)$ vanishing eigenvalues.

We give an example of degeneracy. The Bell state

$$|\Psi_-^{AB}\rangle = \frac{1}{\sqrt{2}}(|0^A, 1^B\rangle - |1^A, 0^B\rangle) \tag{8.20}$$

is, like all other Bell states, entangled. The associated bases of $\mathcal{H}^A$ and $\mathcal{H}^B$ are e.g. $\{|0^A\rangle, |1^A\rangle\}$ and $\{|1^B\rangle, -|0^B\rangle\}$. The state $|\Psi_-^{AB}\rangle$ exhibits spherical symmetry, since one can readily verify that

$$|\Psi_-^{AB}\rangle = \frac{1}{\sqrt{2}}(|0_{\mathbf{r}}^A, 1_{\mathbf{r}}^B\rangle - |1_{\mathbf{r}}^A, 0_{\mathbf{r}}^B\rangle) \tag{8.21}$$

holds, where $|0_{\mathbf{r}}\rangle$ and $|1_{\mathbf{r}}\rangle$ are the eigenvectors of $\mathbf{r}\boldsymbol{\sigma}$ with an arbitrary Bloch vector $\mathbf{r}$ (cf. Eq. (3.31)). This demonstrates that in this case, the Schmidt decomposition is not unique. If several p_n in Eq. (8.9) are the same, the corresponding vectors $|u_n^A, w_n^B\rangle$ can be replaced by linear combinations. This corresponds to the fact that for $|\Psi_-^{AB}\rangle$, the eigenvectors of the reduced density operators

$$\rho^A = \rho^B = \frac{1}{2}\mathbb{1} \tag{8.22}$$

are not determined.

Purification by a transition to a product space In Sect. 9.1.2, we will make use of the following ancillary theorem: *For each system with a density operator ρ^A on $\mathcal{H}^A$, there exists a pure state $|\phi^{AB}\rangle$ in $\mathcal{H}^A \otimes \mathcal{H}^B$ such that ρ^A is the associated reduced density operator:*

$$\rho^A = \mathrm{tr}_B[|\phi^{AB}\rangle\langle\phi^{AB}|] \,. \tag{8.23}$$

For the proof, we consider the ONB $\{|u_n^A\rangle\}$ in which ρ^A is diagonal:

$$\rho^A = \sum_{n=1}^{a} p_n |u_n^A\rangle\langle u_n^A| \,. \tag{8.24}$$

According to the Schmidt decomposition, then

$$|\phi^{AB}\rangle = \sum_{n=1}^{a} \sqrt{p_n}|u_n^A, w_n^B\rangle \tag{8.25}$$

with any ONB $\{|w_n^B\rangle\}$ is a possible purification. Other purifications can be obtained from it by unitary transformations of this basis. This makes it once again clear that there is no unique inverse to taking the partial trace. There are infinitely many composite systems S^{AB} which have the same state of S^A.

8.3.3 The Entropy of the Subsystems as a Measure of Entanglement

Pure states of a bipartite system S^{AB} are either in a product state; then the subsystems are not correlated, or they are entangled, and then they are EPR-correlated (see Sect. 8.1). Correlations can in this case be only non-classical. We denote the entropy of the composite system by $S(AB)$ and the entropies of the subsystems, obtained through the reduced density operators $\rho^A = \mathrm{tr}_B[\rho^{AB}]$ and $\rho^B = \mathrm{tr}_A[\rho^{AB}]$, by $S(A)$ and $S(B)$. Since we consider only pure states, we always have the maximum information about the composite state:

$$S(AB) = 0 \,. \tag{8.26}$$

In a product state, the maximum information about the state of the subsystems is also present:

$$S(A) = S(B) = 0 \,. \tag{8.27}$$

In contrast, for example the density operators of the Bell states are completely without structure

$$\rho^A = \rho^B = \frac{1}{2}\mathbb{1} \,. \tag{8.28}$$

The subsystems in this case are maximally mixed and maximally undetermined:

$$S(A) = S(B) = 1 \,. \tag{8.29}$$

The entropy (indeterminacy of the state) of the subsystems is a measure of the missing information – compared with the pure state of the composite system. If one considers only the subsystems, one loses all the more information about the composite state, the more information is contained in the correlations between the substates. The greater the entropy of the subsystems, the more strongly is the pure state of the composite system correlated and thus entangled. *We therefore take for a pure state $|\psi^{AB}\rangle$ of the composite system the value $E(\psi)$ of the entropy of the subsystems*

$$0 \leq E(\psi) := S(A) = S(B) \leq 1 \tag{8.30}$$

as a measure of the entanglement of the state. $E(\psi)$ is also called the *entropy of entanglement*. With Eqs. (6.4), (6.5), and (8.13), we obtain for the entanglement

$$E(\psi) = -\sum_{n=1}^{k} p_n \ln p_n \tag{8.31}$$

with the Schmidt coefficients p_n. This quantitative measure of the entanglement holds only for pure qubit states. For mixed states, there is a series of suggested measures of the entanglement, which are the same as $E(\psi)$ for the special case of pure states of bipartite systems (see Sect. 8.8). For systems which are composed of two qubits, we give a measure of the entanglement for mixtures in Sect. 11.7.

The entropy of entanglement depends only on the Schmidt coefficients. It is independent of the basis and does not change under local unitary transformations. For $|\psi^{AB'}\rangle = U^A \otimes U^B|\psi^{AB}\rangle$, we have $E(\psi') = E(\psi)$. States in $\mathcal{H}_d \otimes \mathcal{H}_d$, for which $E(\psi) = \log d$ with $d = \dim \mathcal{H}$ applies, are called *maximally entangled*. Bell states are an example.

8.3.4 Subsystems in Pure States are not Entangled*

We prove a theorem for mixtures which we have already mentioned above for pure states of S^{AB}: *If for a bipartite system S^{AB} which is in a mixed state ρ^{AB}, the reduced density operator ρ^A of a subsystem S^A is that of a pure state, then ρ^{AB} is separable.* Since we are seeking a statement about mixtures, we cannot have recourse to the Schmidt decomposition.

We first discuss the case that ρ^A is a mixture. ρ^A is a positive operator. It can be decomposed into its orthonormal eigenstates, which we complete to obtain an ONB $|u_n^A\rangle$

$$\rho^A = \sum_n r_n |u_n^A\rangle\langle u_n^A| \,, \qquad \sum_n r_n = 1 \,, \qquad r_n \geq 0 \,. \tag{8.32}$$

*The sections marked with an asterisk * can be skipped over in a first reading.

Correspondingly, for ρ^{AB}

$$\rho^{AB} = \sum_q s_q |\psi_q^{AB}\rangle\langle\psi_q^{AB}| \,, \qquad \sum_q s_q = 1 \,, \qquad s_q > 0 \,. \tag{8.33}$$

It is expedient for calculations to write the eigenvectors $|\psi_q^{AB}\rangle$ making use of the relative states $|\tilde{w}_n^{(q)B}\rangle$ to $|u_n^A\rangle$ (see Sect. 7.4.1). For clarity, we dispense with the indices A and B and with the tilde so far as possible. We obtain

$$|\psi_q^{AB}\rangle = |\psi_q\rangle = \sum_n |u_n, w_n^{(q)}\rangle \tag{8.34}$$

and thus

$$\begin{aligned} \rho^{AB} &= \sum_{q,n,m} s_q |u_n \,, w_n^{(q)}\rangle\langle u_m, w_m^{(q)}| \\ &= \sum_{n,m} |u_n\rangle\langle u_m| \otimes \sum_q s_q |w_n^{(q)}\rangle\langle w_m^{(q)}| \,. \end{aligned} \tag{8.35}$$

This is in general not a separable state. For the reduced density operator ρ^A of the subsystem S^A, it follows that

$$\rho^A = \mathrm{tr}_B[\rho^{AB}] = \sum_{n,m} |u_n\rangle\langle u_m| \sum_q s_q \langle w_m^{(q)}|w_n^{(q)}\rangle \,. \tag{8.36}$$

To take the trace with the ONB $\{|v_i\rangle\}$ of $\mathcal{H}^B$, we made use of the relation

$$\sum_i \sum_q s_q \langle v_i|w_n^{(q)}\rangle\langle w_m^{(q)}|v_i\rangle = \sum_q s_q \langle w_m^{(q)}|w_n^{(q)}\rangle \,. \tag{8.37}$$

With Eqs. (8.32) and (8.36), we find the matrix elements of ρ^A to be

$$\langle u_k|\rho^A|u_l\rangle = r_k \delta_{kl} = \sum_q s_q \langle w_l^{(q)}|w_k^{(q)}\rangle \quad . \tag{8.38}$$

If, in particular, $r_k = 0$, then we have

$$\sum_q s_q ||w_k^{(q)}|| = 0 \tag{8.39}$$

and, with $s_q > 0$,

$$||w_k^{(q)}|| = 0 \,, \qquad |w_k^{(q)}\rangle = 0 \tag{8.40}$$

for all q.

With this result, we return to the original problem. ρ^A is supposed to be a pure state. Without limitations, we set

$$\rho^A = |u_1\rangle\langle u_1| \,. \tag{8.41}$$

Then, $r_{n\neq 1} = 0$ and thus $|w_{n\neq 1}^{(q)}\rangle = 0$. Inserting into Eq. (8.35) gives

$$\rho^{AB} = |u_1\rangle\langle u_1| \otimes \sum_q s_q |w_1^{(q)}\rangle\langle w_1^{(q)}| \,. \tag{8.42}$$

The density operator ρ^{AB} is thus separable. This yields a remarkable physical result: *If a system is in a pure state, it cannot be entangled with another system.* There are no correlations between this system and any other arbitrary system.

The postulates from Sect. 2.1 refer to pure states. The systems described by these postulates are therefore necessarily also isolated with respect to EPR correlations. We note an additional result: a measurement on a subsystem with a non-degenerate measured value transforms the state of this subsystem into a pure state. *Therefore, this projective measurement breaks up the entanglement with other subsystems, independently of whether the composite system was previously in a mixed or in a pure state.* Once again, a projective measurement is seen to be a strong intervention.

8.4 The PPT Criterion for the Entanglement of Mixtures *

Partial transposition We wish to give a criterion for the entanglement of mixtures. To this end, we introduce the *transposition* T in $\mathcal{H}$. Referring to the computational basis $\{|n\rangle\}$ of $\mathcal{H}$, it is defined as the mapping which transforms $\rho = \sum_{n,m} \rho_{nm}|n\rangle\langle m|$ into

$$T(\rho) = (\rho)^T := \sum_{n,m} \rho_{nm}|m\rangle\langle n|, \quad \rho_{n,m}^T = \rho_{m,n} \,. \tag{8.43}$$

For a self-adjoint matrix, the eigenvalues are the same as those of the transposed matrix. ρ is a positive operator. Therefore, ρ^T is also a positive operator. Transposition is a linear and positive mapping.

Referring to the matrix representation with the ONB $\{|n^A\rangle\}$ and $\{|\mu^B\rangle\}$ of $\mathcal{H}^A$ or $\mathcal{H}^B$, the *partial transposition* in $\mathcal{H}^A$, $(T^A \otimes \mathbb{1}^B)\rho^{AB} =: (\rho^{AB})^{T_A}$ is given by the mapping of ρ^{AB}:

$$\rho^{AB} \leftrightarrow \rho_{m\mu,n\nu} = \langle m^A, \mu^B|\rho^{AB}|n^A, \nu^B\rangle \tag{8.44}$$

onto

$$(\rho^{AB})^{T_A} \leftrightarrow \rho_{m\mu,n\nu}^{T_A} = \rho_{n\mu,m\nu} \,. \tag{8.45}$$

For the partial transposition in $\mathcal{H}^B$, we have correspondingly

$$\langle m^A, \mu^B|(\rho^{AB})^{T_B}|n^A, \nu^B\rangle = \langle m^A, \nu^B|\rho^{AB}|n^A, \mu^B\rangle \,. \tag{8.46}$$

The partial transpositions are positive on all separable density operators (cf. (8.4)), since for example

$$T^A \otimes \mathbb{1}^B(\rho^A \otimes \sigma^B) = (\rho^A)^{T_A} \otimes \sigma^B \,. \tag{8.47}$$

*The sections marked with an asterisk * can be skipped over in a first reading.

It is important that for Hilbert spaces of low dimension, the converse also holds. This is the content of a theorem which we shall not prove ([HHH 96], [HHH 01]). It specifies the *PPT criterion* or positive partial transpose criterion for entanglement: *A state ρ^{AB} in $\mathcal{H}_2^A \otimes \mathcal{H}_2^A$ or $\mathcal{H}_2^A \otimes \mathcal{H}_3^A$ is separable if and only if $(\rho^{AB})^{T_A}$ is positive ($(\rho^{AB})^{T_A} \geq 0$).* This criterion is also called the *Peres-Horodecki criterion*. It is of fundamental significance. Its use in practice, however, is hindered by the fact that one must first determine the state ρ^{AB}.

An example We give an example for the application of the PPT criterion. We consider the Bell state $|\Phi_+^{AB}\rangle$, which is "impurified" by an admixture of the non-entangled maximally-mixed state $\frac{1}{4}\mathbb{1}$

$$\rho^{AB} = p|\Phi_+^{AB}\rangle\langle\Phi_+^{AB}| + (1-p)\frac{1}{4}\mathbb{1}^{AB}, \quad 0 \leq p \leq 1\,. \tag{8.48}$$

We enumerate the computational basis of $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$ by

$$\begin{aligned} |0^A, 0^B\rangle &\leftrightarrow 1\,, \\ |0^A, 1^B\rangle &\leftrightarrow 2\,, \\ |1^A, 0^B\rangle &\leftrightarrow 3 \\ |1^A, 1^B\rangle &\leftrightarrow 4\,.; \end{aligned} \tag{8.49}$$

then we find for the matrix representation

$$\rho^{AB} \leftrightarrow \frac{1}{2}\begin{pmatrix} p+\frac{1}{2}(1-p) & 0 & 0 & p \\ 0 & \frac{1}{2}(1-p) & 0 & 0 \\ 0 & 0 & \frac{1}{2}(1-p) & 0 \\ p & 0 & 0 & p+\frac{1}{2}(1-p) \end{pmatrix}. \tag{8.50}$$

For simplicity, we carry out the transposition T_B in $\mathcal{H}_2^B$. We can read off from Eq. (8.46) that it has the same effect on the matrix elements as the permutation $(1 \to 2, 2 \to 1, 3 \to 4, 4 \to 3)$ of the indices. One thus obtains the partially transposed matrix by transposing each of the four submatrices

$$\rho^{AB} \leftrightarrow \begin{pmatrix} \begin{pmatrix} & \end{pmatrix} & \begin{pmatrix} & \end{pmatrix} \\ \begin{pmatrix} & \end{pmatrix} & \begin{pmatrix} & \end{pmatrix} \end{pmatrix}. \tag{8.51}$$

In this manner, one can readily verify that in this case the resulting matrix has the eigenvalues

$$\lambda_{1,2,3} = \frac{1}{4}(1+p) \geq 0,\ \lambda_4 = \frac{1}{4}(1-3p)\,. \tag{8.52}$$

The PPT criterion states that ρ^{AB} of Eq. (8.48) is entangled if and only if $p > \frac{1}{3}$. If we take ρ^{AB} to be a statistical mixture, then the probability for the occurrence of the entangled Bell state must be greater than $\frac{1}{3}$ for ρ^{AB} to be entangled. The completely mixed state $\frac{1}{4}\mathbb{1}^{AB}$ (it corresponds to $p = 0$) is not entangled.

The general criterion The PPT criterion is a special case of a general criterion which applies to Hilbert spaces $\mathcal{H}_d$ of arbitrary finite dimensions. We give this general criterion without proof ([HHH 96], [HHH 01]) and thereby anticipate the concepts and physical interpretations in Sect. 14.1, which we do not wish to reproduce here. *We consider positive mappings Λ, which map operators on $\mathcal{H}_d^B$ onto operators on $\mathcal{H}_d^B$. ρ^{AB} is a density operator on $\mathcal{H}_d^A \otimes \mathcal{H}_d^B$. It is separable if and only if*

$$(\mathbb{1} \otimes \Lambda)(\rho^{AB}) \geq 0 \tag{8.53}$$

for all positive, but not completely positive mappings Λ. According to the PPT criterion, it is not necessary to test all the Λ's in low dimensions. The evaluation of all transpositions suffices.

This more general criterion is also a purely mathematical one, which cannot be directly implemented as a verification procedure for separability. In [HE 02], a direct experimental verification of entanglement is described, which requires no *a priori* knowledge about a quantum state.

8.5 The Production of Entangled States

Entanglement as the normal case The production of entangled states from the theoretical point of view is quite simple. Let a composite system, which is composed of two subsystems, be in a product state

$$|\psi^{AB}\rangle = |\phi^A\rangle \otimes |\chi^B\rangle \tag{8.54}$$

at the time t_0. If the system experiences a dynamic evolution for $t > t_0$ with a unitary operator U^{AB}, which is not a product operator,

$$U^{AB} \neq U^A \otimes U^B \ , \tag{8.55}$$

– this is the normal case – then it will evolve into an entangled state. *In this sense, entanglement is the "normal state".* In practice, the production of well-determined entangled states (such as the Bell states) or given types of quantum objects (such as photons) requires an experimental effort. Entanglement is frequently broken by perturbations which arise from the environment as a participating third system. We will discuss an example of this in Sect. 14.4.3.

Cascade photons At present, there are a number of experimental methods for producing entangled states in the laboratory. We will discuss one of these as an example. For the detection of correlations, the polarisations of pairs of photons are particularly suitable, since photons can propagate without major perturbations over typical laboratory distances and even much further. We describe a source of entanglement which is operational in the optical range. More details can be found in the references of Sect. 8.8.

An atom decays via two sequential transitions in a *cascade* from an excited state via an intermediate state to its ground state. Such a cascade can be observed e.g. using calcium atoms (see Fig. 8.2). The two photons which are emitted have the wavelengths $\lambda_A = 551,3$ nm and

$\lambda_B = 422,7$ nm. They are in general not emitted in opposite directions. In the experiment, however, one chooses pairs in which e.g. the photon with the wavelength λ_A is emitted in the positive z-direction and the photon of wavelength λ_B in the negative direction. Since in the $J = 0 \to J = 1 \to J = 0$ transition, the total angular momentum J remains unchanged, these photons must be circularly polarised and have equal but opposite angular momenta. It follows furthermore from the details of the atomic decay that the composite state of the two photons has even parity. These two conditions must be obeyed by the state which describes the two-photon system.

As we have seen in Sect. 3.6, the right-hand and left-hand circular polarised states with a fixed propagation vector $\mathbf{k}$ form a basis in the Hilbert space $\mathcal{H}_2$ of polarisations. For the photon pair, $|R^A, R^B\rangle, |R^A, L^B\rangle, |L^A, R^B\rangle, |L^A, L^B\rangle$ is thus a basis of the product space $\mathcal{H}^A \otimes \mathcal{H}^B$; A and B denote here the positive and the negative z-directions, respectively, in which the two photons are emitted. The wavevector $\mathbf{k}$ is in the one case proportional to $\mathbf{e}_z$ and in the other to $-\mathbf{e}_z$. Due to the vanishing total angular momentum, the two-photon state can only be a linear combination of $|R^A, R^B\rangle$ and $|L^A, L^B\rangle$ (compare Fig. 8.3)

$$|\phi^{AB}\rangle = \alpha|R^A, R^B\rangle + \beta|L^A, L^B\rangle\,. \tag{8.56}$$

The physics of the atomic transitions dictates an additional condition: the two-photon state must have even parity. This means that when the coordinate system is changed from right-handed to left-handed, the state remains invariant. This is clearly possible only for a particular choice of α and β:

$$|\Phi_+^{AB}\rangle = \frac{1}{\sqrt{2}}\left(|R^A, R^B\rangle + |L^A, L^B\rangle\right)\,. \tag{8.57}$$

The two-photon state is thus a Bell state, due to the symmetries of the production process, which it must reflect.

The entanglement becomes plausible if one considers the fact that the intermediate state with $J = 1$ is degenerate. The ground state can therefore be arrived at via two different intermediate states (cf. Fig. 8.2). Both "paths" are possible. In analogy to the "paths" in a two-slit experiment, the resulting states interfere with each other. Thereby on the one "path" two right-hand circularly-polarised photons are emitted, and on the other, two photons with left-hand circular polarisation. This correlation leads, together with the superposition, to entanglement.

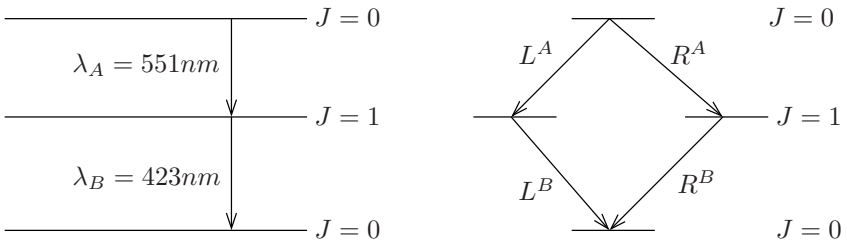


Figure 8.2: Transition scheme of cascade photons.

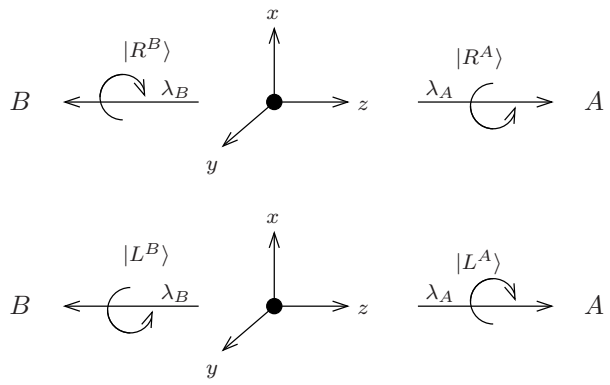


Figure 8.3: Polarisations of the cascade photons

The Bell state $|\Phi_+^{AB}\rangle$ of Eq. (8.57) reflects precisely the situation of the photon pairs which results from their production process.

For later reference, we compute the linear polarisations in the x-direction and the y-direction. To do this, we can make use of Eqs. (3.65) and (3.66); however, we must keep in mind the convention in Fig. 8.3 for the direction of propagation. We find

$$\begin{aligned} |R^A\rangle &= \frac{1}{\sqrt{2}}(|x^A\rangle + i|y^A\rangle) \\ |L^A\rangle &= \frac{1}{\sqrt{2}}(|x^A\rangle - i|y^A\rangle) \\ |R^B\rangle &= \frac{1}{\sqrt{2}}(|x^B\rangle - i|y^B\rangle) \\ |L^B\rangle &= \frac{1}{\sqrt{2}}(|x^B\rangle + i|y^B\rangle) \end{aligned} \tag{8.58}$$

and thus for the entangled state

$$|\Phi_+^{AB}\rangle = \frac{1}{\sqrt{2}}(|x^A, x^B\rangle + |y^A, y^B\rangle)\,. \tag{8.59}$$

In our computation, the position of the x-axis and of the y-axis was not fixed. Equation (8.59) holds for arbitrary orientations. The state $|\Phi_+^{AB}\rangle$ is rotationally symmetric with respect to the z-axis. We shall return to this two-photon state in Chap. 10.

8.6 The No-Cloning Theorem Prevents Transfer of Information Faster than the Velocity of Light

We have seen that a measurement on subsystem A instantaneously transforms subsystem B into a well-defined state. The word "instantaneous" is seductive. Let us imagine that an entangled pure state has been produced whose subsystem S^A is at Alice's location and whose

other subsystem S^B at Bob's location, very far away. Alice attempts to transmit one bit of information to Bob by measuring one of two non-commuting observables on her subsystem S^A. If Bob succeeds in reading out this information on his subsystem S^B, then it would have been transmitted at a velocity greater than that of light and this would contradict the theory of relativity.

We describe in detail the procedure suggested. Alice and Bob each have access to a photon from a polarisation-entangled pair which is in e.g. the Bell state

$$|\Psi_-\rangle = \frac{1}{\sqrt{2}}(|H,V\rangle - |V,H\rangle) = \frac{1}{\sqrt{2}}(|-45^\circ, 45^\circ\rangle + |45^\circ, -45^\circ\rangle) \; . \tag{8.60}$$

Alice carries out her measurement either in the $\{|H\rangle, |V\rangle\}$ basis or in the basis which is rotated by 45° with the vectors $|+45^\circ\rangle = (|H\rangle + |V\rangle)/\sqrt{2}$ and $|-45^\circ\rangle = (|H\rangle - |V\rangle)/\sqrt{2}$. These two basis systems are known to Bob. The choice of the one or the other basis by Alice is the information which is to be transmitted.

If Alice places her photon via a measurement in a particular state (e. g. $|-45^\circ\rangle$), the photon at Bob's location enters the state which is perpendicular (in this example $|+45^\circ\rangle$). This means more precisely that an analyser with this orientation will register a count with certainty. If Bob could measure the polarisation of his photon, then he could read Alice's message. However, for his measurement, Bob can only choose one of the two bases randomly, and he has only one photon on which to carry out the measurement. If Bob's photon as in our example is in the state $|+45^\circ\rangle$ and he measures in the basis $\{|H^B\rangle, |V^B\rangle\}$, then either his detector for H-polarisation or his detector for V-polarisation can register a count. The results of measurements on pairs of photons which were prepared in the state $|\psi_-\rangle$ are correlated this way. A single measurement therefore does not suffice for a certain determination of the $(+45^\circ)$-polarisation. Bob can thus not read out the information. If he however had a machine which could produce many copies of his photon, then he could measure average values. With them, he could determine the state and the transmission of information would be possible (compare Sect. 4.4). This raises the question as to whether quantum systems can be cloned.

The no-cloning theorem We want to prove that there is no machine which can copy arbitrary unknown pure quantum states. We first describe a different situation, in which copying would be possible. We show that there is a suitable copier for orthogonal states $|0\rangle$ and $|1\rangle$. This is the corresponding CNOT gate from Sect. 7.8.1. If the control qubit has the form $|\psi^A\rangle = \alpha|0\rangle + \beta|1\rangle$ and $|0^B\rangle$ is chosen as target qubit, then the action of the CNOT gate in the basis $\{|0\rangle, |1\rangle\}$ consists of the entanglement

$$|\psi^A, 0^B\rangle \to |\phi^{AB}\rangle = \alpha|0,0\rangle + \beta|1,1\rangle \; . \tag{8.61}$$

The two orthogonal states of the control qubit, $|0\rangle$ or $|1\rangle$ (i. e. $\beta = 0$ or $\alpha = 0$), are thus copied by the gate which is adapted to this basis. Except however for these two limiting cases $\alpha = 0$ and $\beta = 0$, which is the trivial case, the superposition $|\psi^A\rangle$ will not be copied in this manner, since copying would have to lead to the product state $|\psi^A\rangle|\psi^B\rangle$.

We now turn to the general case. A quantum system S^A is in the state $|\psi^A\rangle$. This state is supposed to be copied, in other words a second quantum system S^B, which is originally in

the state $|i^B\rangle$, is supposed to be transformed into the state $|\psi^B\rangle$[4]. Here, the initial state $|i^C\rangle$ of the copier system S^C, which completes the composite system, can itself be transformed into a new state $|f(\psi)^C\rangle$ in a manner which depends upon $|\psi^B\rangle$. The overall procedure is supposed to be universal, as with an office copying machine; i.e. a copy of an arbitrary state of S^A is supposed to be made using the same unitary transformation U of the composite system.

$|\varphi^A\rangle$ is a second state which is to be copied. Then the requirement is:

$$|\psi^A\rangle|i^B\rangle|i^C\rangle \quad \xrightarrow{U} \quad |\psi^A\rangle|\psi^B\rangle|f(\psi)^C\rangle \tag{8.62}$$

$$|\varphi^A\rangle|i^B\rangle|i^C\rangle \quad \xrightarrow{U} \quad |\varphi^A\rangle|\varphi^B\rangle|f(\varphi)^C\rangle\,. \tag{8.63}$$

The unitary transformation maintains the inner product

$$\langle\psi^A|\varphi^A\rangle = \langle\psi^A|\varphi^A\rangle\langle\psi^B|\varphi^B\rangle\langle f(\psi)^C|f(\varphi)^C\rangle\,. \tag{8.64}$$

If $|\psi^A\rangle$ and $|\varphi^A\rangle$ are not orthogonal, ($\langle\psi^A|\varphi^A\rangle \neq 0$), it follows that

$$1 = \langle\psi^B|\varphi^B\rangle\langle f(\psi)^C|f(\varphi)^C\rangle\,. \tag{8.65}$$

We consider the absolute values. Since all the states are normalised, with $|\langle\psi^B|\varphi^B\rangle| \leq 1$ and $|\langle f(\psi)^C|f(\varphi)^C\rangle| \leq 1$, then

$$|\langle\psi^B|\varphi^B\rangle| = 1 \quad \text{i.e.} \quad |\psi^A\rangle = |\varphi^A\rangle \tag{8.66}$$

is a necessary condition for fulfilling (8.65). Therefore, the machine could not copy a second state $|\varphi^A\rangle$ which is not orthogonal to $|\psi^A\rangle$. *There is no universal copier for pure quantum states (no-cloning theorem).*

Thus, the attempt described above to construct a contradiction between the theory of relativity and quantum theory must fail. The conflict-free coexistence of the two theories is remarkable, since the requirements of relativity were not taken into account in the formulation of quantum mechanics in its nonrelativistic form, which we are using.

8.7 Marking States by Entanglement*

8.7.1 Which-Way Marking*

Two-path interferometer We remind the reader of the Mach-Zehnder interferometer treated in Sect. 3.7.2 (see Fig. (3.8)), which is traversed by a single quantum object. *Behind* the first beam splitter, we have the state ρ. The quantum system in this state encounters a second beam splitter with a phase shifter (cf. Fig. 3.10). A detector at the 0 output can then register an interference pattern as a function of the set phase shift α. More precisely, the signal probability $p(\alpha)$ will be a function of α. According to Eq. (3.89), for an arbitrary incident state ρ, it is found to be

$$p_\rho(\alpha) = \text{tr}[\rho|\alpha\rangle\langle\alpha|] \tag{8.67}$$

[4] We speak for simplicity (see Sect. 2.1.2) again in an abbreviated manner of transforming states into other states. More precisely, the copier functions in such a manner that it carries out – controlled by S^A – the same preparation procedure for S^B that S^A had originally passed through.

[*] The sections marked with an asterisk * can be skipped over in a first reading.

with

$$|\alpha\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\alpha}|1\rangle) \; . \tag{8.68}$$

For the case that the state behind the first beam splitter is a pure state $|\chi\rangle$, then the signal probability and the fringe contrast of the interference pattern are those given in Sect. 3.7.2. We discuss different states ρ.

In the interpretation of the mixture

$$\rho = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2}\mathbb{1} \tag{8.69}$$

as a statistical mixture, the quantum object would be incident on the phase shifter and the second beam splitter either in the state $|0\rangle$ or in $|1\rangle$, each with the relative probability 1/2. In all these cases there is no α dependence of p and thus no interference:

$$p_{|0\rangle}(\alpha) = p_{|1\rangle}(\alpha) = \frac{1}{2}, \quad p_{|0\rangle}(\alpha) + p_{|1\rangle}(\alpha) = p_\rho(\alpha) = 1 \; . \tag{8.70}$$

The fringe contrast ν (cf. Sect. 3.7.2) is

$$\nu_{|0\rangle} = \nu_{|1\rangle} = \nu_\rho = 0 \; . \tag{8.71}$$

If the pure state $|\chi\rangle$ is a superposition of $|0\rangle$ and $|1\rangle$, then it can in general not be associated with a particular mode (a particular path). Correspondingly, interference then occurs. The general qubit vector $|\chi(\Theta,\varphi)\rangle$ and the vector perpendicular to it, $|\chi_\perp(\Theta,\varphi)\rangle$, have the form (compare Eq. (3.90)):

$$|\chi(\Theta,\varphi)\rangle = \cos\frac{\Theta}{2}|0\rangle + e^{i\frac{\varphi}{2}}\sin\frac{\Theta}{2}|1\rangle \tag{8.72}$$

$$\begin{aligned} |\chi_\perp(\Theta,\varphi)\rangle &= -\sin\frac{\Theta}{2}|0\rangle + e^{i\frac{\varphi}{2}}\cos\frac{\Theta}{2}|1\rangle \\ &= |\chi(\pi+\Theta,\varphi)\rangle \; . \end{aligned} \tag{8.73}$$

They make up an ONB. With $p(\alpha)$ from Eq. (3.89), we find for the signal probability of the detector as a function of the phase shift α

$$p_{|\chi\rangle}(\alpha) = \frac{1}{2}[1 + \sin\Theta\cos(\alpha - \frac{\varphi}{2})] \tag{8.74}$$

$$p_{|\chi_\perp\rangle}(\alpha) = \frac{1}{2}[1 - \sin\Theta\cos(\alpha - \frac{\varphi}{2})] \; . \tag{8.75}$$

For the fringe contrast, we obtain

$$\nu_{|\chi\rangle} = \nu_{|\chi_\perp\rangle} = \sin\Theta \; . \tag{8.76}$$

If the pure state behind the first beam splitter is not $|0\rangle$ or $|1\rangle$ ($\theta \neq 0, \pi$), then on variation of α, an interference pattern results with a fringe contrast which depends on Θ. The interference patterns of $|\chi\rangle$ and $|\chi_\perp\rangle$ are shown for $\varphi = \pi$ in Fig. (8.4).

One can see that on addition of the probabilities, the dependence on the phase shift α just cancels. The interference pattern vanishes:

$$p_{|\chi\rangle}(\alpha) + p_{|\chi_\perp\rangle}(\alpha) = p_\rho(\alpha) = 1 \; . \tag{8.77}$$

This is the situation which occurs for the equally-weighted statistical mixture

$$\rho = \frac{1}{2}(|\chi\rangle\langle\chi| + |\chi_\perp\rangle\langle\chi_\perp|) = \frac{1}{2}\mathbb{1} \tag{8.78}$$

of the basis states $|\chi\rangle$ and $|\chi_\perp\rangle$. The mixture effects an addition of the interference patterns which are shifted relative to each other. Interference fringes are no longer visible.

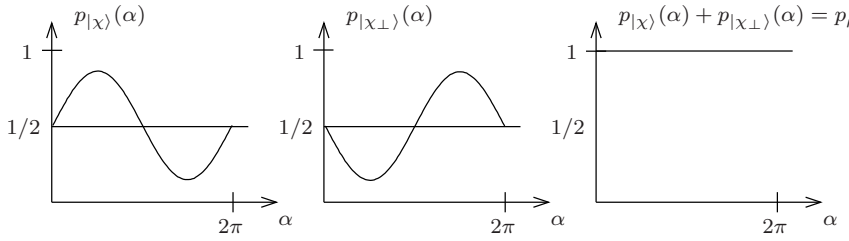


Figure 8.4: The interference patterns for the state $|\chi\rangle$ and the state perpendicular to it, $|\chi_\perp\rangle$.

Entanglement with marker states We consider the example of single atoms in a Mach-Zehnder atom interferometer. The results however are generally valid. The atom, in the qubit state of Eq. (8.72),

$$|\chi^A\rangle = c_0|0^A\rangle + c_1|1^A\rangle \tag{8.79}$$

after the first beam splitter is supposed to interact with an additional qubit system S^M in the interferometer before it reaches the second beam splitter, in such a way that the entangled state

$$|\phi^{AM}\rangle = c_0|0^A, 0^M\rangle + c_1|1^A, 1^M\rangle \tag{8.80}$$

of the composite system S^{AM} is formed. The atomic states $|0^A\rangle$ and $|1^A\rangle$ are "marked" by the associated *marker states* $|0^M\rangle$ or $|1^M\rangle$, which are likewise eigenstates of σ_z. S^M is called the *marker system*. In the present case, we are dealing with *which-way markers*, corresponding to the definitions of the states $|0^A\rangle$ and $|1^A\rangle$. Markers can be e.g. the internal degrees of freedom of an atom or other quantum objects.

The interaction between S^A and S^M, which gives rise to the marking, must effect the following transition:

$$\left(c_0|0^A\rangle + c_1|1^A\rangle\right)|i^M\rangle \rightarrow c_0|0^A, 0^M\rangle + c_1|1^A, 1^M\rangle \; . \tag{8.81}$$

If one leaves questions of the physical implementation out of consideration, it is simple to specify a unitary transformation which has this effect:

$$U^{AB} = |0^A\rangle\langle 0^A| \otimes |0^M\rangle\langle 0^M| + |1^A\rangle\langle 1^A| \otimes |1^M\rangle\langle 1^M| \; . \tag{8.82}$$

We choose as initial state for the marker system $|i^M\rangle = \frac{1}{\sqrt{2}}(0^M\rangle + |1^M\rangle)$. U^{AB} is clearly a non-local transformation. In the special case that the atom enters along the path 0 (i. e. is in the state $|0^A\rangle$), the marker is brought into the "position" $|0^M\rangle$ and correspondingly, into the position $|1^M\rangle$ for path 1. Only when the atom is initially in a superposition does an entangled state result.

Loss of interference capability What physical results are produced by the which-way marking of Eq. (8.80)? We carry out measurements on the marker qubit in the computational basis $\{|0^M\rangle, |1^M\rangle\}$, or, expressed differently, we measure the *marker observable* $|0^M\rangle\langle 0^M| - |1^M\rangle\langle 1^M|$. When the measurement result is $+1$, corresponding to $|0^M\rangle$, the atom continues along the path 0; for -1, along the path 1. *Measurement of the marker observables thus breaks up the interference and determines the path of the atom.*

Remarkably, it is not even necessary for loss of the interference capability that a measurement be carried out on the marker system. After the entangled state $|\phi^{AM}\rangle$ of Eq. (8.80) with the reduced density operator

$$\rho^A = \mathrm{tr}_M[|\phi^{AM}\rangle\langle\phi^{AM}|] = |c_0|^2|0^A\rangle\langle 0^A| + |c_1|^2|1^A\rangle\langle 1^A| \tag{8.83}$$

has been formed by marking, the state ρ^A of the atomic system is the same in terms of all probability statements as a statistical mixture of the states $|0^A\rangle$ and $|1^A\rangle$. Accordingly, the interference vanishes. *When one marks interfering states, the interference capability is lost even if no measurements are carried out on the markers.* The production of correlations in the system S^{AB} destroys the local coherence in the system S^A.

The information which was contained before the marking in the pure state $|\chi^A\rangle$ of the atom and which determined the interference pattern is no longer stored in the atomic state and cannot be read out by making a measurement on the atom. Indeed, it has however not been lost through the unitary entanglement dynamics. It was deposited as mutual information in the correlations with the marker system. We now want to see how one can call it up again by measuring the correlations, and can use them to make suitable selections.

8.7.2 Quantum Erasure*

We consider for simplicity the special case $c_1 = c_2 = \frac{1}{\sqrt{2}}$ (i. e. a Bell state).

$$|\Phi_+^{AM}\rangle = \frac{1}{\sqrt{2}}\left(|0^A, 0^M\rangle + |1^A, 1^M\rangle\right) \tag{8.84}$$

$$\rho^A = \frac{1}{2}\mathbb{1}^A\,. \tag{8.85}$$

The influence of the marking destroys the interference. If the original interference pattern can be reconstructed by a clever procedure, one speaks of *quantum erasure*. The considerations in the previous chapter on the basis states $|\chi(\Theta,\varphi)\rangle$ and $|\chi_\perp(\Theta,\varphi)\rangle$ with $\Theta \neq 0, \pi$ give an indication of how we can extract information from the mixture ρ^A at hand. In our case, ρ^A

*The sections marked with an asterisk * can be skipped over in a first reading.

is not a statistical mixture (blend), but rather the reduced density operator of a subsystem. It cannot however be distinguished by measurements from a statistical mixture of the states $|\chi^A\rangle$ and $|\chi^A_\perp\rangle$. If it were possible after the marking to firstly transform the atoms into the states $|\chi^A\rangle$ or $|\chi^A_\perp\rangle$ and secondly to separate the atoms in the state $|\chi^A\rangle$ from those in the state $|\chi^A_\perp\rangle$ by selection, then the corresponding atomic ensembles would lead to the interference patterns in Fig. 8.4. Since the composite system ρ^{AM} is in a pure state $|\Phi^{AM}_+\rangle$, we can indeed carry out both these tasks by suitable measurements on the marker system S^M.

In order to give rise to quantum erasure in the present case, we proceed as follows: it can be readily shown that the entangled state $|\Phi^{AM}_+\rangle$ of Eq. (8.84) can be written with *rotated marker states*

$$|\Lambda^M(\Theta,\varphi)\rangle = \cos\frac{\Theta}{2}|0^M\rangle + e^{-i\frac{\varphi}{2}}\sin\frac{\Theta}{2}|1^M\rangle \tag{8.86}$$

$$|\Lambda^M_\perp(\Theta,\varphi)\rangle = -\sin\frac{\Theta}{2}|0^M\rangle + e^{-i\frac{\varphi}{2}}\cos\frac{\Theta}{2}|1^M\rangle \tag{8.87}$$

in a form which is completely analogous to Eq. (8.84)

$$|\Phi^{AM}_+\rangle = \frac{1}{\sqrt{2}}(|\chi^A,\Lambda^M\rangle + |\chi^A_\perp,\Lambda^M_\perp\rangle)\,. \tag{8.88}$$

Thus, for the unmodified composite state $|\Phi^{AM}_+\rangle$, the interference-capable states $|\chi^A\rangle$ and $|\chi^A_\perp\rangle$ instead of the states $|0^A\rangle$ and $|1^A\rangle$ are entangled with the new marker states. A measurement on the marker system in the rotated basis $\{|\Lambda^M\rangle, |\Lambda^M_\perp\rangle\}$, with measured values $+1$ or -1, then transforms the atomic state into $|\chi^A\rangle$ if the measured value $+1$ is registered, or into $|\chi^A_\perp\rangle$ for the measured value -1. In a non-selective measurement, the resulting reduced density operator ρ^A of Eq. (8.78) however still does not yield an interference pattern.

The decisive second step thus consists of a *selection* after the measurement, i.e. of an *unmixing of the mixture*. This reverses the addition in Eq. (8.77) and in Fig. 8.4. If we now consider only the contributions to the interference pattern of those atoms which were prepared by the measurement with the measured value $+1$, then we obtain the interference pattern of Fig. 8.4 for $p_{|\chi\rangle}(\alpha)$ from Eq. (8.74). Correspondingly, after selection for the measured value -1, the interference curve $p_{|\chi_\perp\rangle}(\alpha)$ from Eq. (8.75) is obtained. *A selective measurement on the marker system S^M in a suitably rotated marker basis reproduces the interference pattern for the atomic systems S^A.* The fringe contrast $\nu = 1$ is attained for $\Theta = \frac{\pi}{2}$. For $\varphi = 0$, the marker states are the eigenstates of σ_x. Measurement in rotated bases with subsequent selection is a frequently-used method (see e.g. Sects. 10.1 and 15.2.2).

8.7.3 Delayed Choice of the Marker Observables*

A thought experiment Without claiming that it can be readily implemented, we wish to imagine the following experimental setup in a thought experiment: an atom is within an atomic interferometer of the Mach-Zehnder type. In the state $|0^A\rangle$, it emits a photon with vertical polarisation $|0^M\rangle$; in the state $|1^A\rangle$, the photon is emitted with horizontal polarisation $|1^M\rangle$. The photon serves as a marker system S^M. Its two polarisation states are the marker states. The state of the composite system is again supposed to be $|\Phi^{AM}_+\rangle$.

*The sections marked with an asterisk * can be skipped over in a first reading.

We have seen already in Sect. 7.5 that for the correlations between the measured results, it is unimportant in which order the measurements on the subsystems are carried out. That is true here, also; we thus need not carry out the measurements on the marker system first. One obtains the same results by registering the detector signals from all the atoms which arrive and enumerating the results. The associated photons are likewise numbered in the same fashion in this thought experiment and initially stored separately without measurement. Only in a later phase of the experiment are the photons then individually measured with the rotated marker observables (rotated analyser directions $|\Lambda^M\rangle$ and $|\Lambda^M_\perp\rangle$) and the results labelled with the photon numbers. If the measured results on the atoms and the photons are finally correlated in terms of their numbers, then one obtains from the measurements on the atoms – following the selection described above – the interference pattern in Fig. 8.4. The results of the measurements on the atoms alone contain no information. Only the selected data, i.e. the subensemble, yields the interference effect.

After the atoms have passed through the interferometer and have already been detected, one can however by a choice of the marker observables determine whether an interference pattern will by obtained from the atom measurements after selection, $\big(|\Lambda^M\rangle, |\Lambda^M_\perp\rangle\big)$; or no interference pattern results, $\big(|0^M\rangle, |1^M\rangle\big)$. This could be understood to imply that one can still choose between the alternatives (i) atoms arrive "via only one of the paths" ("particle behaviour" and thus no interference pattern) and (ii) atoms arrive in the sense of a superposition "via both paths" ("wave behaviour" and thus an interference pattern), even after the measurements have been carried out on all the atoms. This *delayed choice* appears to imply an influence on the behaviour of the system (as "wave" or "particle") in the past. What is false in this interpretation?

The assumption is false that for an atom, between preparation and measurement, one can speak of an event "on one path" or of a "wave" (compare Sect. 2.1.4). The system of atoms S^A is in a completely mixed state before the measurement. The alternatives (i) and (ii) in the form (i) "no interference pattern" and (ii) "interference pattern" are not already realised at the time of the measurements on the atoms or even before, but rather only after carrying out the rotated measurement on the photons and the selection of the measurement results on the atoms. The selection however has no effect on the past. Before the selection, there was nothing present which could be associated with the alternatives (i) or (ii) for the atoms. *There is thus no "delayed-choice" paradox in the case of delayed choice of the marker observables* (see also Sect. 11.5).

8.8 Complementary Topics and Further Reading

- Production of entanglement: [Aul 00], [BEZ 00], [NC 00], [BZ 02], [DM 02], [Hei 02], [SS 04].
- The concept of classically-correlated quantum states: [Wer 89].
- Separability criteria for systems with two or more subsystems, measures of entanglement for mixtures: [Wer 89], [LBC 00], [HHH 01], [Ter 02], [Bru 02], [Cir 02], [DHR 02], [Key 02].

- Discussion of the PPT criterion in a wider context: [HHH 01], [Cir 02].
- The no-cloning theorem. The classic article: [WZ 82]. Review article: [SIG 05]
- The no-cloning theorem states that there are no procedures with which cloning can be carried out with certainty. Correspondingly, the proof refers to unitary and thus to deterministic evolution. If one allows additional selective measurements, then cloning becomes possible. However, copies are produced with only a certain probability (*probabilistic cloning*): [DG 98].
- Quantum erasure: [Eng 99], [ESW 99].
- Which-way experiments, marking in experiments: [DNR 98], [Eng 99], [ESW 99], [DR 00], [Rem 06].
- The "delayed choice" discussion and experiments: [Whe 78], [HWZ 87], [HKW 95], [Aul 00, Chap. 26].

8.9 Problems for Chapter 8

Prob. 8.1 [for 8.1.1]: Prove the assertion from Sect. 8.1.1 following Eq. (8.1) about the characterisation of correlated states.

Prob. 8.2 [for 8.1.2]: Prove that the separability condition (8.3) is equivalent to the requirement

$$\rho^{AB} = \sum_l q_l |\phi_l^A\rangle\langle\phi_l^A| \otimes |\chi_l^B\rangle\langle\chi_l^B| \tag{8.89}$$

with $0 \leq q_l \leq 1$ and $\sum_l q_l = 1$. Comparison with the generalised operator Z^{AB} from Eq. (7.20) shows that the requirement (8.89) indeed implies a restriction.

Prob. 8.3 [for 8.3]: Show that for the state

$$|\phi^{ABC}\rangle = \frac{1}{\sqrt{2}} |0^A\rangle(|0^B, 0^C\rangle + |1^B, 1^C\rangle) \,, \tag{8.90}$$

there is no Schmidt decomposition

$$|\phi^{ABC}\rangle = \sum_n \sqrt{p_n} |u_n^A\rangle|v_n^B\rangle|w_n^C\rangle \tag{8.91}$$

with orthonormal Schmidt bases $\{|u_n^A\rangle\}, \{|v_n^B\rangle\}, \{|w_n^C\rangle\}$.

Prob. 8.4 [for 8.4]: Verify Eq. (8.58).

Prob. 8.5 [for 8.7]: How would the considerations about quantum erasure be modified if one started with the state (8.80) instead of the state (8.84)?

Prob. 8.6 [for 8.7]: Prove Eq. (8.88). Does a form invariance of this type also hold for the other Bell states?

© 2007 WILEY-VCH Verlag GmbH & Co.

9 Correlations and Non-Local Measurements

Quantum information can be stored in the correlations between subsystems. We wish to make this assertion more precise. We base our discussion on the considerations of Chap. 6 on quantum entropy. This concept leads us again to a renewed understanding of entanglement which complements the results obtained in Chap. 7. In composite systems, above and beyond the measurements on subsystems, non-local measurements are also possible, which allow us among other things to read out information which is stored non-locally.

9.1 Entropies and the Correlations of Composite Quantum Systems

9.1.1 Mutual Information as a Measure of Correlations

Additivity and subadditivity We begin by deriving a useful relation. Making use of the decomposition of two density operators ρ^A and ρ^B in the ONB of $\mathcal{H}^A$ and $\mathcal{H}^B$,

$$\begin{aligned} \rho^A &= \sum_n a_n |n^A\rangle\langle n^A|, \\ \rho^B &= \sum_j b_j |j^B\rangle\langle j^B| \,. \end{aligned} \tag{9.1}$$

we obtain for the logarithm of the product state (cf. Eq. (1.49))

$$\begin{aligned} \log\left(\rho^A \otimes \rho^B\right) &= \sum_{n,j} \log(a_n b_j) |n^A, j^B\rangle\langle n^A, j^B| \\ &= (\log \rho^A)\mathbb{1}^B + \mathbb{1}^A(\log \rho^B) \,. \end{aligned} \tag{9.2}$$

If ρ^A and ρ^B are the reduced density operators of a density operator ρ^{AB}, then with Eqs. (9.2) and (7.38) we find

$$\begin{aligned} \mathrm{tr}_{AB}[\rho^{AB} \log(\rho^A \otimes \rho^B)] &= \mathrm{tr}_{AB}[\rho^{AB}(\log \rho^{AB})\mathbb{1}^B] - \mathrm{tr}_{AB}[\rho^{AB}\mathbb{1}^A(\log \rho^B)] \\ &= \mathrm{tr}_A[\rho^A \log \rho^A] - \mathrm{tr}_B[\rho^B \log \rho^B] \,. \end{aligned} \tag{9.3}$$

With the definition of the *joint entropy* for composite systems,

$$S(AB) := S(A,B) := S(\rho^{AB}) := -tr_{AB}[\rho^{AB} \log \rho^{AB}] \,, \tag{9.4}$$

the *additivity* of the entropy of the subsystems

$$S(\rho^A \otimes \rho^B) = S(\rho^A) + S(\rho^B) \tag{9.5}$$

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

follows for a product state $\rho^{AB} = \rho^A \otimes \rho^B$. If ρ^{AB} is not a product state, we can derive only an estimate, instead of Eq. (9.5). We use Klein's inequality (6.24) as a starting point, and apply it to the product space $\mathcal{H}^A \otimes \mathcal{H}^B$. In Eq. (6.17), we replace ρ by ρ^{AB} and σ by $\sigma^{AB} = \rho^A \otimes \rho^B$ with the reduced density operators ρ^A and ρ^B of ρ^{AB}. We then find by applying Eq. (9.3):

$$S(\rho^{AB}) \leq -\mathrm{tr}_{AB}[\rho^{AB} \log(\rho^A \otimes \rho^B)] = S(\rho^A) + S(\rho^B) \,. \tag{9.6}$$

We write the result in the form

$$S(AB) \leq S(A) + S(B) \,. \tag{9.7}$$

This property is termed the *subadditivity* of the entropy of composite quantum systems. The equals sign holds if and only if the subsystems S^A and S^B are independent of each other: $\rho^{AB} = \rho^A \otimes \rho^B$. The analogous equation for classical systems is Eq. (5.43) (see Fig. 5.3). *If the subsystems are not independent of each other, and the composite state ρ^{AB} is thus not separable, then the system as a whole contains more information than can be read out of all the subsystems together.* We will elucidate this with an example.

Mutual information of the subsystems In the classical as in the quantum-mechanical case, the additional information lies in the correlations between the systems. In order to describe this quantitatively for composite quantum systems, we introduce the *mutual information* $S(A:B)$ *as a measure of the degree of correlation of the subsystems*, in analogy to Eq. (5.31):

$$S(A:B) := S(A) + S(B) - S(AB) \geq 0 \,. \tag{9.8}$$

$S(A:B)$ indicates for a state ρ^{AB} just how much the uncertainty of the composite system, expressed by its entropy S^{AB}, is less than that of the subsystems, S^A and S^B together (cf. Eq. (9.7)). Or, differently formulated: $S(A:B)$ *is a measure of how much more information is stored in the composite system than in the subsystems.* $S(A:B)$ *can at the same time characterise the distance of the state ρ^{AB} from the non-entangled state $\rho^A \otimes \rho^B$.*

9.1.2 The Triangle Inequality

Let the system S^{AB} be in a state ρ^{AB}. We have seen in Sect. 8.3.2 that this state can always be purified. This means that one can always add a system S^C to S^{AB} and then find a pure state in the enlarged composite system S^{ABC}, such that the reduced density operator of the subsystem S^{AB} is just ρ^{AB}.

We apply the inequality for subadditivity:

$$S(C) + S(A) \geq S(AC) \,. \tag{9.9}$$

Since the system S^{ABC} is in a pure state, the reduced density operators are the same on decomposition into two subsystems. We have already demonstrated this in connection with the Schmidt decomposition.

$$S(AC) = S(B), \qquad S(C) = S(AB) \,. \tag{9.10}$$

Inserting into Eq. (9.9) leads to

$$S(AB) \geq S(B) - S(A) \; . \tag{9.11}$$

The systems S^A and S^B enter the relation symmetrically. Thus,

$$S(AB) \geq S(A) - S(B) \tag{9.12}$$

also holds, and therefore

$$S(AB) \geq |S(A) - S(B)| \; . \tag{9.13}$$

This is the *triangle inequality*, which is sometimes also called the *Araki-Lieb inequality*.

We have shown in Eq. (5.39) that for Shannon's entropy of classical systems, the relation

$$H(A,B) \geq \left\{ \begin{array}{c} H(A) \\ H(B) \end{array} \right\} \tag{9.14}$$

is always fulfilled. The indeterminacy of the composite system exceeds that of every individual system. This cannot hold for quantum systems. The Bell states, for which we showed that $S(AB) = 0$ and $S(A) = S(B) = 1$, are a straightforward counterexample.

9.1.3 Entangled *vs.* Classically-Correlated Quantum Systems

We want to compare the entropy of the subsystems S^A and S^B, which is supposed to be equal to one in all cases ($S(A) = S(B) = 1$), with the entropy of the composite system by using the example of a bipartite system S^{AB} in three different states. The subsystems are taken for simplicity to be qubits. Our goal is to relate correlations and entanglement to the entropy, and in particular to the mutual information $S(A:B)$, making use of Eq. (9.8). The situation is indicated graphically in Fig. 9.1.

Example I: independent subsystems In the first state, the subsystems are completely independent of one another, i.e. ρ^{AB} is a product state. We choose

$$\rho^{AB} = \frac{1}{4}\mathbb{1}^{AB} = \left(\frac{1}{2}\mathbb{1}^A\right) \otimes \left(\frac{1}{2}\mathbb{1}^B\right) = \rho^A \otimes \rho^B \; . \tag{9.15}$$

There are no correlations between the subsystems. The quantum entropies are found immediately to be

$$\begin{aligned} S(AB) &= \log 4 = 2, & (9.16)\\ S(A) &= S(B) = \log 2 = 1 & (9.17)\\ S(A:B) &= 0 \; . & (9.18) \end{aligned}$$

For later comparison, we also note ρ^{AB} in the computational basis of $\mathcal{H}^A \otimes \mathcal{H}^B$:

$$\begin{aligned} \rho^{AB} &= \frac{1}{4}(|0^A,0^B\rangle\langle 0^A,0^B| + |0^A,1^B\rangle\langle 0^A,1^B| \\ &\quad + |1^A,0^B\rangle\langle 1^A,0^B| + |1^A,1^B\rangle\langle 1^A,1^B|) \; . \end{aligned} \tag{9.19}$$

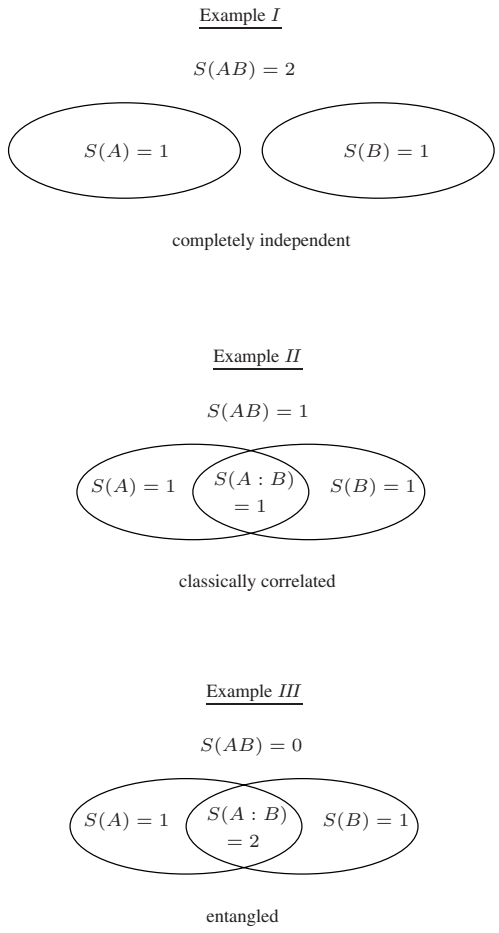


Figure 9.1: The mutual information $S(A : B)$ for the same entropy of the subsystems.

Example II: classically-correlated subsystems We want to establish correlations without creating entanglement. We can do this for example with the *separable mixture*

$$\rho^{AB} = \frac{1}{2}(|0^A\rangle\langle 0^A| \otimes |0^B\rangle\langle 0^B| + |1^A\rangle\langle 1^A| \otimes |1^B\rangle\langle 1^B|) \tag{9.20}$$

$$= \frac{1}{2}(|0^A, 0^B\rangle\langle 0^A, 0^B| + |1^A, 1^B\rangle\langle 1^A, 1^B|) \tag{9.21}$$

of pure product states. ρ^{AB} contains fewer terms in comparison to the ρ^{AB} of Eq. (9.19). In the computational basis of $\mathcal{H}^A \otimes \mathcal{H}^B$, the density operator ρ^{AB} has the matrix representation

$$\rho^{AB} = \text{diag}\left(\frac{1}{2}, 0, 0, \frac{1}{2}\right) \ . \tag{9.22}$$

Only diagonal terms occur. The entropy is then found to be

$$S(AB) = -2\left[\frac{1}{2}\log\frac{1}{2}\right] = 1\ . \tag{9.23}$$

The reduced density operators of the subsystems are the same as those in the other examples:

$$\rho^A = \frac{1}{2}\mathbb{1}^A\ , \qquad \rho^B = \frac{1}{2}\mathbb{1}^B\ . \tag{9.24}$$

From this we obtain again

$$S(A) = S(B) = 1 \tag{9.25}$$

but in this case a non-vanishing mutual information,

$$S(A:B) = 1\ . \tag{9.26}$$

It can be seen immediately from Eq. (9.21) that the results of measurements in the computational basis on the two subsystems are correlated: if the measurement on S^A yields the measured value belonging to $|0\rangle$, then the measurement on S^B also gives this measured value; the corresponding results hold for $|1\rangle$. The subsystems are termed *classically correlated*, since they were prepared by LOCC and since the measurements on one subsystem does not modify the well-defined (correlated) pure state of the other subsystem. This means in other words that the state of S^{AB} is a separable mixture.

Example III: entangled subsystems Here, we take as a simple example the Bell state

$$|\Phi_+^{AB}\rangle = \frac{1}{\sqrt{2}}(|0^A, 0^B\rangle + |1^A, 1^B\rangle)\ . \tag{9.27}$$

Since it is a pure state, we have

$$S(AB) = 0\ . \tag{9.28}$$

As in the two previous examples, we find

$$\rho^A = \frac{1}{2}\mathbb{1}^A, \qquad \rho^B = \frac{1}{2}\mathbb{1}^B \tag{9.29}$$

and thus

$$S(A) = S(B) = 1\ . \tag{9.30}$$

For a Bell state, the mutual information takes on the largest possible value

$$S(A:B) = 2\ . \tag{9.31}$$

Comparison of the three examples The fact that in the case of entanglement, the results of measurements on the subsystems are correlated, has already been discussed. In connection with Bell's inequality, we will show explicitly in Chap. 10 that the correlations in examples II and III are quantitatively different. Here, we want to describe this difference by making use of the entropy.

In all three cases, owing to $S(A) = S(B) = 1$, we have the same indeterminacy of the states of the subsystems. The indeterminacy $S(AB)$ of the state of the overall system is, in contrast, different in each case (cf. Fig. 9.1). The *mutual information* $S(A:B)$ specifies, due to

$$S(AB) = S(A) + S(B) - S(A:B)\ , \tag{9.32}$$

the extent to which the actual entropy of the composite system is smaller than the total entropy of the independent subsystems (separable composite state). Less entropy means less indeterminacy of the state. In all three examples, the states of the subsystems are maximally undetermined. Nevertheless, in example III, the entangled composite state is maximally determined. The information which had been lacking for this is localised completely in the correlations and can be visualised through $S(A:B)$. In example II, the correlations are not sufficient to completely determine the state of S^{AB} as a pure state. Correspondingly, the overall system is in a mixed state. $S(A:B)$ is in this case smaller. *As we have seen, the mutual quantum entropy $S(A:B)$ refers not only to entanglement, but also reflects the classical correlations.* The goal of reading out the information which is stored in the correlations leads to non-local measurements.

9.2 Non-Local Measurements

9.2.1 The Bell Basis

We first take note of some properties of the Bell states

$$|\Phi_{\pm}^{AB}\rangle = \frac{1}{\sqrt{2}}(|0^A,0^B\rangle \pm |1^A,1^B\rangle) \tag{9.33}$$

$$|\Psi_{\pm}^{AB}\rangle = \frac{1}{\sqrt{2}}(|0^A,1^B\rangle \pm |1^A,0^B\rangle)\ . \tag{9.34}$$

They form an orthonormal basis (*Bell basis*) of the product space $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$ and are maximally entangled. The reduced density operators are maximally mixed, $\rho^A = \rho^B = \frac{1}{2}\mathbb{1}$. *By intervention on only one subsystem (i.e. locally without classical communication), one can by using the σ operators in a unitary fashion transform a Bell state into any other Bell state.* We give an example:

$$\sigma_1^A:\ |\Psi_+^{AB}\rangle \to \frac{1}{\sqrt{2}}(|1^A,1^B\rangle + |0^A,0^B\rangle) = |\Phi_+^{AB}\rangle \tag{9.35}$$

$$\sigma_2^A:\ |\Psi_+^{AB}\rangle \to \frac{-i}{\sqrt{2}}(|0^A,0^B\rangle - |1^A,1^B\rangle) = -i|\Phi_-^{AB}\rangle \tag{9.36}$$

$$\sigma_3^A:\ |\Psi_+^{AB}\rangle \to \frac{1}{\sqrt{2}}(|0^A,1^B\rangle - |1^A,0^B\rangle) = |\Psi_-^{AB}\rangle\ . \tag{9.37}$$

Frequently in computations, products of Pauli operators occur, whose matrix elements are to be evaluated in the Bell basis (e.g. $\langle\Phi_+^{AB}|\sigma_1^A\sigma_3^B|\Phi_+^{AB}\rangle$). In such cases, it is expedient to relate the actions of the Pauli operators in $\mathcal{H}_2^B$ to those in $\mathcal{H}_2^A$; then all of the relations for the Pauli operators derived in Sect. 3.1 (e. g. Eq. (3.11)) can be directly applied. We give an example to which we shall also return later:

$$\sigma_1^A: \ |\Phi_+^{AB}\rangle \longrightarrow |\Psi_+^{AB}\rangle\,, \qquad \sigma_1^B: \ |\Phi_+^{AB}\rangle \longrightarrow |\Psi_+^{AB}\rangle \tag{9.38}$$

$$\sigma_2^A: \ |\Phi_+^{AB}\rangle \longrightarrow -i|\Psi_-^{AB}\rangle\,, \qquad \sigma_2^B: \ |\Phi_+^{AB}\rangle \longrightarrow i|\Psi_-^{AB}\rangle \tag{9.39}$$

$$\sigma_3^A: \ |\Phi_+^{AB}\rangle \longrightarrow |\Phi_-^{AB}\rangle\,, \qquad \sigma_3^B: \ |\Phi_+^{AB}\rangle \longrightarrow |\Phi_-^{AB}\rangle\,. \tag{9.40}$$

The action of σ_3^B on $|\Phi_+^{AB}\rangle$ can be replaced by the action of σ_3^A, etc. There are corresponding relations for all the vectors of the Bell basis. One can for example confirm in this way that

$$\begin{aligned}\langle\Phi_+^{AB}|\sigma_1^A\sigma_3^B|\Phi_+^{AB}\rangle &= \langle\Phi_+^{AB}|\sigma_1^A\sigma_3^A|\Phi_+^{AB}\rangle \\ &\sim \langle\Phi_+^{AB}|\sigma_2^A|\Phi_+^{AB}\rangle \sim \langle\Phi_+^{AB}|\Psi_-^{AB}\rangle = 0\end{aligned} \tag{9.41}$$

holds.

We note an additional mathematical property of the Bell states. They are eigenvectors of products of the σ operators:

$$\sigma_1^A\sigma_1^B|\Phi_\pm^{AB}\rangle = \pm|\Phi_\pm^{AB}\rangle \tag{9.42}$$

$$\sigma_1^A\sigma_1^B|\Psi_\pm^{AB}\rangle = \pm|\Psi_\pm^{AB}\rangle \tag{9.43}$$

$$\sigma_2^A\sigma_2^B|\Phi_\pm^{AB}\rangle = \mp|\Phi_\pm^{AB}\rangle \tag{9.44}$$

$$\sigma_2^A\sigma_2^B|\Psi_\pm^{AB}\rangle = \pm|\Psi_\pm^{AB}\rangle \tag{9.45}$$

$$\sigma_3^A\sigma_3^B|\Phi_\pm^{AB}\rangle = +|\Phi_\pm^{AB}\rangle \tag{9.46}$$

$$\sigma_3^A\sigma_3^B|\Psi_\pm^{AB}\rangle = -|\Psi_\pm^{AB}\rangle\,. \tag{9.47}$$

9.2.2 Local and Non-Local Measurements

As we have seen in Sect. 7.3.1, non-locality can occur in three ways:

(i) States obtained from a non-local preparation procedure can be non-separable.

(ii) There are non-local unitary transformations.

(iii) We saw in Chap. 7 that measurements on composite systems can be described by Hermitian operators in the product space $\mathcal{H}^A \otimes \mathcal{H}^B \otimes \mathcal{H}^C \otimes \ldots$. Their eigenvalues are the possible measured values. By selective measurements, the corresponding properties can be prepared. There are measurements whose measured values and the resulting states cannot be obtained by carrying out measurements separately on each individual subsystem. Instead, one requires an apparatus which carries out measurements on two or more subsystems together. These measurements, to which we turn now, are called *non-local measurements* (they are also termed global or joint or collective measurements). They measure the values of *non-local observables*.

We again consider a bipartite system S^{AB}. *An Hermitian observable G^{AB} on $\mathcal{H}^A \otimes \mathcal{H}^B$ is in general non-local*. For clarity, we will limit ourselves initially to the special case that G^{AB} is a product operator $C^A \otimes D^B$ on $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$. It describes the non-local measurement of the collective physical quantity $C^A D^B$ (to be read as a single symbol). As an illustration, we consider the observable $\sigma_3^A \sigma_3^B$ of a bipartite system. As Eqs. (9.46) and (9.47) show, the associated measured values $+1$ and -1 are degenerate. The result of the measurement which yields the value $+1$ is that the state is projected onto the subspace spanned by $|\Phi_+^{AB}\rangle$ and $|\Phi_-^{AB}\rangle$. All the Bell states remain unchanged after measurement of the observables $\sigma_3^A \, \sigma_3^B$. The degeneracy makes itself known in the fact that along with $|\Phi_\pm^{AB}\rangle$ and $|\Psi_\pm^{AB}\rangle$, the vectors of the computational basis of $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$,

$$\begin{array}{rclcrcl} \sigma_3^A \sigma_3^B |0,0\rangle & = & |0,0\rangle, & \qquad & \sigma_3^A \sigma_3^B |1,1\rangle & = & |1,1\rangle \\ \sigma_3^A \sigma_3^B |0,1\rangle & = & -|0,1\rangle, & & \sigma_3^A \sigma_3^B |1,0\rangle & = & -|1,0\rangle \end{array} \tag{9.48}$$

are likewise eigenstates with the eigenvalues $+1$ or -1.

The difference between the global measurement $\sigma_3^A \sigma_3^B$ and two local measurements of σ_3 (i.e. a measurement of $\sigma_3^A \otimes \mathbb{1}^B$ and $\mathbb{1}^A \otimes \sigma_3^B$) becomes clear if one considers the states into which the system is transformed by these measurements. If Bob measures the observable σ_3^B (the spin in z-direction) in a local fashion on the subsystem B of a 2-qubit system in the state $|\Phi_+^{AB}\rangle$ and the measurement yields the result -1, then the system will be left in the state $|1^A, 1^B\rangle$. The initial state $|\Phi_+^{AB}\rangle$ is no longer present. The entanglement is broken by the measurement. A subsequent measurement of σ_3^A by Alice on the subsystem S^A yields the result -1, and leaves the state $|1^A, 1^B\rangle$ unchanged as the final state. In contrast, a global measurement of $\sigma_3^A \sigma_3^B$ leads to the final state $|\Phi_+^{AB}\rangle$. This example demonstrates that *mathematically, the action of the operator $\sigma_3^A \, \sigma_3^B = (\sigma_3^A \otimes \mathbb{1}^B)(\mathbb{1}^A \otimes \sigma_3^B)$ comes about through* separate *actions of the operators σ_3^A and σ_3^B in $\mathcal{H}_2^A$ or in $\mathcal{H}_2^B$. Physically, however, the measurement of the collective observable $\sigma_3^A \, \sigma_3^B$ is in general* not *equivalent to two local spin measurements $\sigma_3^A \otimes \mathbb{1}^B$ und $\mathbb{1}^A \otimes \sigma_3^B$ by Alice or Bob.*[1] This is verified by the different final states obtained in the two cases.

This result is not limited to product vectors. In general, it holds that *observables require a non-local measurement when there are entangled eigenstates of the observable operator, since local measurements would destroy this entanglement.* This leaves us with a problem, since we are primarily able to carry out local measurements. They are much more readily implemented experimentally. How can we reduce global measurements to a local measuring procedure? We discuss two approaches.

Mean values are locally measurable We first introduce a notation. The index LL (local-local) indicates e.g. in $(C^A \otimes D^B)_{LL}$ that the measurements are to be carried out locally by measurement of the commuting observables $C^A \otimes \mathbb{1}^B$ and $\mathbb{1}^A \otimes D^B$. The resulting value is the product of the locally-obtained measurement results. In Sect. 7.5.1, we showed that for

[1] In the second case, projective measurements are carried out e.g. sequentially on S^A and S^B. The resulting state and the product of the measured values can be interpreted as the result of a single generalised measurement. We discuss generalised measurements in Chaps. 13 and 16. In this connection, an interesting direct comparison with global measurements is also possible.

the mean values (expectation values), the relation

$$\mathrm{tr}_{AB}[(C^A \otimes D^B)\rho^{AB}] = \mathrm{tr}_{AB}[(C^A \otimes D^B)_{LL}\rho^{AB}] \tag{9.49}$$

holds. *Mean values of product operators can be determined through local measurements on the subsystems.*

We illustrate this using the example of non-local observables for 2-qubit systems. We have seen in Sect. 3.1 that the operators $\{\sigma_0 := \frac{1}{\sqrt{2}}\mathbb{1}, \frac{1}{\sqrt{2}}\sigma_i; i = 1, 2, 3\}$ form an orthonormal operator basis in the Liouville space $\mathbb{L}$ of the operators on $\mathcal{H}_2$. Correspondingly, the operators $\{\frac{1}{2}\sigma_k^A \otimes \sigma_l^B; k, l = 0, \ldots, 3\}$ are an orthonormal basis in the product Liouville space $\mathbb{L}^{AB} = \mathbb{L}^A \otimes \mathbb{L}^B$ of the operators on $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$. Each operator on $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$ can thus be decomposed in the form

$$G^{AB} = \sum_{k,l=0}^{3} q_{kl}\sigma_k^A \otimes \sigma_l^B \ . \tag{9.50}$$

Therefore, with Eq. (9.49) for the mean value of the observable G^{AB}, a decomposition into local measurements

$$\mathrm{tr}_{AB}[G^{AB}\rho^{AB}] = \sum_{k,l=0}^{3} q_{kl}\mathrm{tr}_{AB}[(\sigma_k^A \otimes \sigma_l^B)_{LL}\rho^{AB}], \quad q_{kl} \in \mathbb{R} \ . \tag{9.51}$$

is obtained. *Mean values can be obtained for every observable of a 2-qubit system by means of local measurements of the observables σ_k, $k = 0, \ldots, 3$ (i.e. in principle by simple local polarisation measurements) on the subsystems.* For this purpose, the mean values of the products of the measured values (i.e. $\mathrm{tr}_{AB}[(\sigma_i^A \otimes \sigma_i^B)_{LL}\rho^{AB}]$) are evaluated using the q_{kl} as in Eq. (9.51). A corresponding conclusion holds for the probabilities of the measured results of non-local measurements, which can be written as mean values of the associated projection operators. These statements are not limited to $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$. We shall return to the discussion of non-local observables in connection with entanglement witnesses in Sect. 10.6.

9.2.3 Non-local Measurements by Means of Local Measurements on an Ancillary System

This method is mainly of theoretical interest. We discuss as the simplest example once again the observable $\sigma_3^A\sigma_3^B$. An arbitrary state $|\chi^{AB}\rangle$ from $\mathcal{H}^{AB}$ can be expanded in terms of the computational basis of $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$:

$$|\chi^{AB}\rangle = c_1|0^A, 0^B\rangle + c_2|1^A, 1^B\rangle + c_3|1^A, 0^B\rangle + c_4|0^A, 1^B\rangle \tag{9.52}$$

($\sum_i |c_i|^2 = 1$). A selective measurement of $\sigma_3^A\sigma_3^B$ (compare (9.48)) transforms the state $|\chi^{AB}\rangle$ – depending on the result, $+1$ or -1, of the measurement – with the probabilities $p(+1)$ or $p(-1)$ into the state $|\chi'^{AB}_+\rangle$ or $|\chi'^{AB}_-\rangle$, respectively:

$$\begin{aligned} +1: \quad |\tilde{\chi}'^{AB}_+\rangle &= c_1|0^A, 0^B\rangle + c_2|1^A, 1^B\rangle \\ p(+1) &= |c_1|^2 + |c_2|^2 \end{aligned} \tag{9.53}$$

$$\begin{aligned} -1: \quad |\tilde{\chi}'^{AB}_-\rangle &= c_3|1^A, 0^B\rangle + c_4|0^A, 1^B\rangle \\ p(-1) &= |c_3|^2 + |c_4|^2 \ . \end{aligned} \tag{9.54}$$

The non-normalised vector $|\tilde{\chi}'^{AB}_{+}\rangle$ is the projection of $|\chi^{AB}\rangle$ onto the subspace spanned by $|0^A, 0^B\rangle$ and $|1^A, 1^B\rangle$ in $\mathcal{H}^{AB}$ with the degenerate eigenvalue $+1$ of $\sigma_3^A\sigma_3^B$. A corresponding statement holds for $|\tilde{\chi}'^{AB}_{-}\rangle$. The probabilities are given by the magnitudes of the non-normalised vectors.

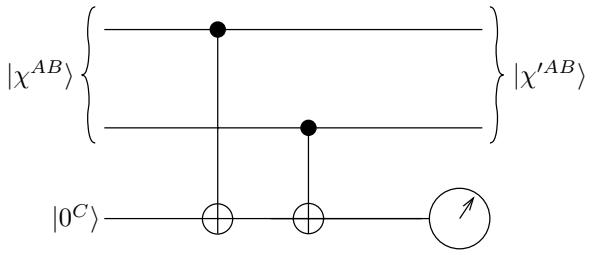


Figure 9.2: After it passes through the quantum circuit, a projective measurement on the ancillary system S^C yields a non-local measurement of $\sigma_3^A\sigma_3^B$ on the system S^{AB}.

This result can also be obtained in the following way by means of a local measurement on an enlarged entangled system. In our special case we implement it by adding an ancillary system $\mathcal{H}_2^C$ to the system S^{AB}. The composite system S^{ABC} passes through the quantum circuit shown in Fig. 9.2. It transforms the initial state $|\chi^{AB}\rangle|0^C\rangle$ by a unitary non-local transformation into the state

$$U|\chi^{AB}\rangle|0^C\rangle = c_1|0^A, 0^B, 0^C\rangle + c_2|1^A, 1^B, 0^C\rangle + c_3|1^A, 0^B, 1^C\rangle + c_4|0^A, 1^B, 1^C\rangle \; . \tag{9.55}$$

A subsequent local projective measurement of σ_3^C on the ancillary system S^C yields the measured values $+1$ and -1 with the probabilities $p(+1)$ and $p(-1)$ from Eqs. (9.53) and (9.54). The corresponding final states of S^{AB} are $|\chi'^{\,AB}_{+}\rangle$ or $|\chi'^{\,AB}_{-}\rangle$. A dynamic implementation of the quantum circuit presupposes that the subsystems S^A and S^B are spatially not too far removed from each other.

We have, with the help of an entangling unitary transformation U on $\mathcal{H}_2^A \otimes \mathcal{H}_2^B \otimes \mathcal{H}_2^C$ and projective measurement on the ancillary system S^C, effected a non-local measurement on the system S^{AB}. The results of the measurement and the probabilities with which they occur can be read off directly from the results of the projective measurements on S^C. This example of a non-local measurement in $\mathcal{H}^{AB}$ reflects a quite general structure for the implementation of generalised quantum measurements. We will illustrate such generalised measurements in more detail in Chap. 13.

With the measurements of the observables $\sigma_3^A\sigma_3^B$ as described, we cannot however distinguish between two states with the same eigenvalue, $+1$ or -1. To do this, one measurement is not sufficient. We describe in the following section how a non-local measurement in the Bell basis can be reduced to two local measurements.

9.2.4 Non-locally Stored Information and Bell Measurements

We refer to Sect. 9.2.1. The measured values $+1$ and -1 of the observables σ_3, which belong to the eigenstates $|0\rangle$ and $|1\rangle$, will be abbreviated by $+$ or $-$ (compare Eq. (3.15)). In the computational basis of $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$, two bits can be stored in the form of the four possibilities $(+,+)$, $(+,-)$, $(-,+)$, $(-,-)$ and read out again by means of two local measurements with $\sigma_3^A \otimes \mathbb{1}^B$ or $\mathbb{1}^A \otimes \sigma_3^B$. The associated projection operators are:

$$P_{++} := |0^A, 0^B\rangle\langle 0^A, 0^B|, \qquad P_{+-} := |0^A, 1^B\rangle\langle 0^A, 1^B| \tag{9.56}$$

$$P_{-+} := |1^A, 0^B\rangle\langle 1^A, 0^B|, \qquad P_{--} := |1^A, 1^B\rangle\langle 1^A, 1^B| \,. \tag{9.57}$$

The information is stored locally. In the Bell basis of $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$, two bits can likewise be stored. We denote them as the *parity bit* Φ or Ψ (i.e. parallel or antiparallel "spins") and as the *phase bit* (of sign $+$ or $-$). As we have already seen, this information is stored as mutual information in the correlations, and it is therefore non-locally stored.

If we measure the observables $\sigma_3^A \otimes \mathbb{1}^B$ and $\mathbb{1}^A \otimes \sigma_3^B$ locally on a system in the Bell state $|\Phi_\pm^{AB}\rangle$ and take the product of the results, we find $+1$. For $|\Psi_\pm^{AB}\rangle$, this procedure correspondingly yields -1. We have thus determined the parity bit locally. Thereafter, however, the initial state is no longer available, but instead a state of the computational basis. We can thus read out only one bit locally. The corresponding conclusion holes for a local measurement of the phase bit by means of $\sigma_1^A \otimes \mathbb{1}^B$ and $\mathbb{1}^A \otimes \sigma_1^B$. *Through local measurements, the complete information stored in Bell states cannot be read out*. We require a non-degenerate projection measurement which projects onto the states of the Bell basis rather than onto the computational basis. The two basis systems can be transformed into one another by means of a unitary transformation in the space $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$. We can make use of this fact.

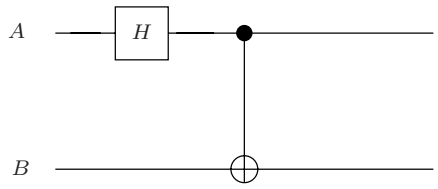


Figure 9.3: A quantum circuit for the production of Bell states.

We note the effect of a Hadamard transformation followed by a CNOT transformation on the computational basis. The quantum circuit for this purpose is shown in Fig. 9.3.

$$\begin{aligned} |0^A, 0^B\rangle \quad &\xrightarrow{H} \quad \frac{1}{\sqrt{2}}(|0^A\rangle + |1^A\rangle)|0^B\rangle \\ &\xrightarrow{\text{CNOT}} \frac{1}{\sqrt{2}}(|0^A, 0^B\rangle + |1^A, 1^B\rangle = |\Phi_+^{AB}\rangle \end{aligned} \tag{9.58}$$

$$|0^A, 1^B\rangle \quad \xrightarrow{H} \quad \frac{1}{\sqrt{2}}(|0^A\rangle + |1^A\rangle)|1^B\rangle \xrightarrow{\text{CNOT}} |\Psi_+^{AB}\rangle \tag{9.59}$$

$$|1^A, 0^B\rangle \xrightarrow{H} \frac{1}{\sqrt{2}}(|0^A\rangle - |1^A\rangle)|0^B\rangle \xrightarrow{\text{CNOT}} |\Phi_-^{AB}\rangle \tag{9.60}$$

$$|1^A, 1^B\rangle \xrightarrow{H} \frac{1}{\sqrt{2}}(|0^A\rangle - |1^A\rangle)|1^B\rangle \xrightarrow{\text{CNOT}} |\Psi_-^{AB}\rangle \; . \tag{9.61}$$

Owing to

$$H = H^\dagger, \quad H^2 = 1 \tag{9.62}$$

$$\text{CNOT} = (\text{CNOT})^\dagger, \quad (\text{CNOT})^2 = 1 \; , \tag{9.63}$$

for the overall unitary transformation applied, we have

$$U^{AB} := (\text{CNOT}) \cdot H, \quad (U^{AB})^{-1} = H \cdot (\text{CNOT}) \; . \tag{9.64}$$

If we allow Bell states to pass through the quantum circuit in the reverse direction, then they are converted by the effect of $(U^{AB})^{-1}$ once again into the corresponding states of the computational basis. The projectors of the two basis systems are related by

$$|\phi_+^{AB}\rangle\langle\phi_+^{AB}| = U^{AB} P_{++}^{AB} (U^{AB})^{-1} \tag{9.65}$$

$$|\psi_+^{AB}\rangle\langle\psi_+^{AB}| = U^{AB} P_{+-}^{AB} (U^{AB})^{-1} \tag{9.66}$$

$$|\phi_-^{AB}\rangle\langle\phi_-^{AB}| = U^{AB} P_{-+}^{AB} (U^{AB})^{-1} \tag{9.67}$$

$$|\psi_-^{AB}\rangle\langle\psi_-^{AB}| = U^{AB} P_{--}^{AB} (U^{AB})^{-1} \; . \tag{9.68}$$

A *Bell measurement* on a state $|\chi^{AB}\rangle$ consists of first subjecting the state to a unitary and non-local transformation with $(U^{AB})^{-1} = H \cdot (\text{CNOT})$ and then carrying out a local measurement on the resulting state in the computational basis. Then one obtains for example the resulting pair of measured values $(+,+)$ with the probability

$$p_{++} = \langle\chi^{AB}|\Phi_+^{AB}\rangle\langle\Phi_+^{AB}|\chi^{AB}\rangle \; . \tag{9.69}$$

It belongs to the projection operator $|\Phi_+^{AB}\rangle\langle\Phi_+^{AB}|$. Correspondingly, with the probability $p_{+-} = \langle\chi^{AB}|\Psi_+^{AB}\rangle\langle\Psi_+^{AB}|\chi^{AB}\rangle$, the pair $(+,-)$ is obtained, etc. Here, $p_{++} + p_{+-} + p_{-+} + p_{--} = 1$ holds. In the special case $|\chi^{AB}\rangle = |\Phi_+^{AB}\rangle$, we find $p_{++} = 1$. Corresponding results apply to the other Bell states. If one codes 2-bit information in Bell states, then this information can be read out unambiguously via a Bell measurement with the correspondences $(+,+) \leftrightarrow |\Phi_+^{AB}\rangle$, $(+,-) \leftrightarrow |\Psi_+^{AB}\rangle$, $(-,+) \leftrightarrow |\Phi_-^{AB}\rangle$, and $(-,-) \leftrightarrow |\Psi_-^{AB}\rangle$.

In order to guarantee that the Bell measurement leads to a Bell state, the state obtained from the selective local measurements in the computational basis must be sent in reverse through the quantum circuit (application of U^{AB}). For e.g. the pair of measured values $(+,+)$, this then prepares the corresponding Bell state $|\Phi_+^{AB}\rangle$.

9.3 Complementary Topics and Further Reading

- Quantum non-locality without entanglement: in [BDF 99], orthogonal non-entangled states of 2-part and 3-part systems are given, which cannot be reliably distinguished by local measurements and classical communication. It is shown that this is however possible by means of non-local measurements. See also [WSH 00] and [WH 02] for discussions of this problem. Holism in quantum mechanics thus expresses itself not only in the entanglement of states.
- Much of what we can readily write down in theory is difficult to implement in practice. It is not possible to carry out a perfect Bell measurement with linear optical elements (no-go theorem): [LCS 99]. Two Bell states can be distinguished in this way, the other two cannot: [BEZ 00], Chap. 3.5.

9.4 Problems for Chapter 9

Prob. 9.1 [for 9.2.2]: Carry out a number of local measurements of $\sigma_3^A \otimes \mathbb{1}^B$ and $\mathbb{1}^A \otimes \sigma_3^B$ on $|\chi^{AB}\rangle$ from Eq. (9.52) one after another. Determine the resulting states and the probabilities for the different pairs of measured values. Take products of the measured values and the corresponding probabilities. Compare with the results of (9.53) and (9.54), as well as with the statement about the expectation values from Sect. 7.5.1.

© 2007 WILEY-VCH Verlag GmbH & Co.

10 There is no (Local-Realistic) Alternative to Quantum Theory

In the previous chapter, we saw that for entangled states, in contrast to the states of classical systems, information is stored non-locally. In the following sections, we describe an approach which allows us to demonstrate in a direct manner that entanglement has no counterpart in classical physics. We wish to show that entanglement is one of the central non-classical structural elements of quantum theory. To this end, we describe as examples two experiments whose results cannot be explained classically.

What sort of experiments will these be? We recall the conjuring trick in Sect. 7.7. We have an immediate experimental access to composite systems via local measurements on the subsystems. If we compare the pairs of measured results, we find typical differences between systems in entangled pure states on the one hand, and non-entangled as well as classical systems on the other. This can serve as a further theoretical characterisation of entanglement. Of greater importance in this connection, however, is the demarcation of quantum mechanics and classical physics. When classical composite systems and entangled quantum systems differ in their correlations and experiments verify the existence of quantum correlations, then we are dealing with a phenomenon which cannot be classically substantiated. *We would thus have demonstrated that quantum mechanics cannot be reduced to classical physics.* This is our goal. In order to carry out this programme, we will describe typical correlation experiments and analyse them both quantum-mechanically and classically. We must however first answer the question, "What does classical mean?" and derive experimentally verifiable consequences. We start by sketching a special experimental setup for photons and computing the correlations which result from the quantum-theoretical calculation. In the next step, we ask whether a classical model can exist which explains these correlations.

10.1 EPR Experiments and Their Quantum-mechanical Explanation

Photons We describe an experiment with the cascade photons which we have already encountered in Sect. 8.5. A source emits photon pairs with different frequencies ν_A and ν_B in opposite directions. They propagate along the z-axis towards the observers at the locations A and B (cf. Fig. 10.1). The distances of the observers from the source may be very great, and they need not be the same. The pairs of photons have entangled polarisation states. They are

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

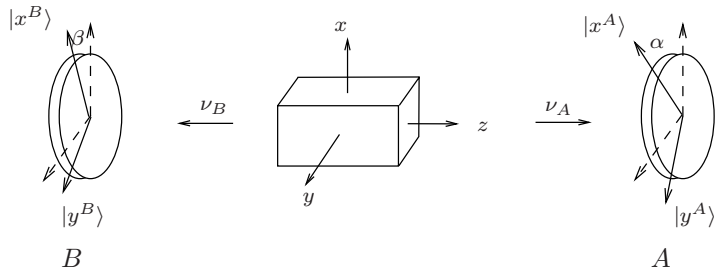


Figure 10.1: A polarisation measurement on pairs of photons. The dashed lines indicate the x- and the y-axes. The solid lines are the rotated axes.

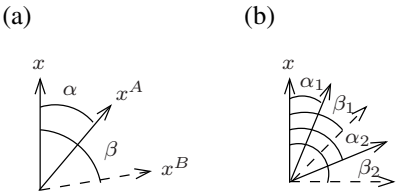


Figure 10.2: The orientations of the analysers.

in the rotationally-symmetric Bell state

$$|\Phi_+^{AB}\rangle = \frac{1}{\sqrt{2}}\left(|x,x\rangle + |y,y\rangle\right) . \tag{10.1}$$

$|x\rangle$ and $|y\rangle$ are states of linear polarisation in the x- and the y-directions, respectively.

The observers detect the linear polarisations of the two photons (compare Fig. 10.2a). To this end, at A, an analyser with the orientation (x^A, y^A) is set up. It is rotated by an angle α around the z-axis with respect to the orientation defined by the x- and y-axes. Behind the analyser, there are two detectors which register counts when the polarisation $|x^A\rangle$ or $|y^A\rangle$ is found, respectively. The polarisation $|x^A\rangle$ is attributed to the measured value $+1$ and the polarisation $|y^A\rangle$ to the value -1. The measurement at B of the second photon is carried out in the same manner. In particular, we want to allow the B analyser to be rotated by an angle $\beta \neq \alpha$ relative to the orientation defined by the coordinate axes. The corresponding polarisation directions are $|x^B\rangle$ and $|y^B\rangle$. We again associate them with the measured values $+1$ or -1, respectively. The observable operators for the local measurements of the photons at A and B are thus

$$E^A(\alpha) = |x^A\rangle\langle x^A| - |y^A\rangle\langle y^A| \tag{10.2}$$

$$E^B(\beta) = |x^B\rangle\langle x^B| - |y^B\rangle\langle y^B| . \tag{10.3}$$

We consider initially measurements only on photons at A or on photons at B. The reduced density operators of the Bell state $|\Phi_+^{AB}\rangle$ are each maximally mixed. The probability of finding the result $+1$ or -1 at A is therefore – independently of the rotation angles α and β – in each case $\frac{1}{2}$. The same is true of the measurement on the second photon at B.

In the next step, we investigate correlations and determine the probabilities for the occurrence of pairs of measured values. We denote by P_{+-} the probability that the x^A polarisation is observed at A and the y^B polarisation at B. This corresponds to the local measured values $+1$ and -1 and to a product of measured values equal to -1. The probabilities for the remaining combinations are denoted correspondingly. For the vectorial polarisation states, we find

$$|x^A\rangle = \cos\alpha|x\rangle + \sin\alpha|y\rangle \tag{10.4}$$

$$|y^A\rangle = -\sin\alpha|x\rangle + \cos\alpha|y\rangle \ . \tag{10.5}$$

Analogous relations hold for $|x^B\rangle$ and $|y^B\rangle$ with the rotation angle β. We thus find for the probability P_{++}

$$P_{++} = \langle\Phi_+^{AB}|x^A, x^B\rangle\langle x^A, x^B|\Phi_+^{AB}\rangle = \frac{1}{2}\cos^2(\beta-\alpha) \tag{10.6}$$

and correspondingly for the other probabilities:

$$P_{--} = P_{++} = \frac{1}{2}\cos^2(\beta-\alpha) \tag{10.7}$$

$$P_{+-} = P_{-+} = \frac{1}{2}\sin^2(\beta-\alpha) \ . \tag{10.8}$$

The fact that only the difference between the rotation angles α and β plays a role is due to the rotational symmetry of the initial state $|\Phi_+^{AB}\rangle$.

The correlation coefficient The measurements at the locations A and B always yield the results 0 or 1. In a later step, we will attempt to give a classical explanation for the results. For comparison, we require a characteristic quantity which can be calculated for both cases. We employ the degree of correlation of the pairs of measured values. It can be characterised solely with reference to the observed experimental results, i.e. independently of the theoretical rationale employed. We take the product of the local measured values at A and B. In the case of complete correlation, we would always obtain $+1$. If for a measured value of e.g. $+1$ at A we now and then obtain not $+1$ at B, but rather -1 (and therefore a product of -1), then the degree of correlation is smaller. This is reflected in the value of the *correlation coefficient* ϵ^{AB}, which is defined as the mean value of the product of the local measured values at A and B.

EPR correlations of entangled photons We first compute the correlation coefficient ϵ^{AB} for the entangled photons. In the special case of parallel orientations of the polarisers ($\alpha = \beta$), we find a complete correlation of the results: if the $|x^A\rangle$ polarisation is found at A, then with certainty the $|x^B\rangle$ polarisation will likewise be observed at B ($P_{+-} = 0$). The same holds for the y polarisations ($P_{-+} = 0$). The two polarisation directions occur at A and B in a completely random fashion with the probabilities $\frac{1}{2}$ in each case. We now turn to the general case:

As we saw in Sect. 7.5.1, the expectation value of the product of the local measured values at A and B can be computed as the mean value of the product operator $E^A \otimes E^B$:

$$\epsilon^{AB}(\alpha,\beta) := \langle\Phi_+^{AB}|E^A(\alpha)\otimes E^B(\beta)|\Phi_+^{AB}\rangle \ . \tag{10.9}$$

Explicit evaluation using Eqs. (10.2) and (10.3) leads to the probabilities of Eqs. (10.7) and (10.8) (compare Eq. (10.6)).

$$\epsilon^{AB}(\alpha,\beta) = P_{++} + P_{--} - P_{+-} - P_{-+} \,. \tag{10.10}$$

Equation (10.10) can also be read out directly. It contains the products of the measured values $+1$ and -1, multiplied by the probabilities for their occurrence. Equation (10.10) is evaluated with Eqs. (10.7) and (10.8) and yields the result

$$\epsilon^{AB}(\alpha,\beta) = \cos 2(\beta-\alpha) \,. \tag{10.11}$$

In the special case of parallel orientation of the analysers ($\alpha = \beta$), the resulting perfect correlation gives $\epsilon^{AB}(\alpha,\beta) = 1$.

Objects with spin-$\frac{1}{2}$ For polarisation measurements on two objects with spin-$\frac{1}{2}$, we can proceed in an analogous manner. For a given $\boldsymbol{e_3}$ direction, the Bell state

$$|\Phi_+^{AB}\rangle = \frac{1}{\sqrt{2}}(|0^A,0^B\rangle + |1^A,1^B\rangle) \tag{10.12}$$

is prepared by a source. $|0\rangle$ and $|1\rangle$ are the eigenstates of σ_3. At A and B, polarisations are measured along the directions $\boldsymbol{a}$ and $\boldsymbol{b}$. In order to simplify the computations, we assume that $\boldsymbol{a}$ and $\boldsymbol{b}$ are perpendicular to $\boldsymbol{e_2}$

$$\boldsymbol{a} = (\sin\alpha, 0, \cos\alpha), \quad \boldsymbol{b} = (\sin\beta, 0, \cos\beta) \,. \tag{10.13}$$

The corresponding observables are

$$E^A(\alpha) = \boldsymbol{\sigma}^A\boldsymbol{a} \,, \qquad E^B(\beta) = \boldsymbol{\sigma}^B\boldsymbol{b} \,. \tag{10.14}$$

We can determine the correlation coefficient ϵ^{AB} of Eq. (10.9) in a simple way by referring to some auxiliary relations which have already been derived. We first apply Eqs. (9.38) and (9.40) and take into account the relation $a_2 = b_2 = 0$. We then make use of Eq. (3.11), apply $\boldsymbol{a}\times\boldsymbol{b} \sim \boldsymbol{e}_2$ as well as Eq. (9.39), and use the orthonormalisation of the Bell vectors. We then find successively for the correlation coefficient ϵ^{AB}:

$$\begin{aligned} \epsilon^{AB}(\alpha,\beta) &= \langle\Phi_+^{AB}|(\boldsymbol{\sigma}^A\boldsymbol{a})(\boldsymbol{\sigma}^B\boldsymbol{b})|\Phi_+^{AB}\rangle = \langle\Phi_+^{AB}|(\boldsymbol{\sigma}^A\boldsymbol{a})(\boldsymbol{\sigma}^A\boldsymbol{b})|\Phi_+^{AB}\rangle \\ &= \langle\Phi_+^{AB}|(\boldsymbol{ab})\mathbb{1}^A|\Phi_+^{AB}\rangle + i\langle\Phi_+^{AB}|\boldsymbol{\sigma}^A(\boldsymbol{a}\times\boldsymbol{b})|\Phi_+^{AB}\rangle \\ &= \boldsymbol{ab} = \cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\beta-\alpha) \,. \end{aligned} \tag{10.15}$$

Compared to ϵ^{AB} for photons, in the case of spinors, the angles are typically half as large.

10.2 Correlated Gloves

We have seen that for a parallel orientation of the analysers at A and B, complete correlation is observed. If the x polarisation is found for a photon at the location A, then with certainty the photon registered at location B will also exhibit the parallel x polarisation and *vice versa*. The same is true of the y polarisation. It is not important whether the measurement is first carried out at A or at B. Furthermore, the distance between A and B can be arbitrarily great (assuming disturbance-free transmission of the photons). Is this a result which holds only for quantum objects?

The pair of gloves We first take up again the discussion of Section 7.5.2 from a different point of view. We take a pair of gloves to consist as usual of a left-hand and a right-hand glove. Each of the gloves is placed into a box. One box is brought to the location A and the other to B. The observers at A and B know that the boxes contain a pair of gloves. If, on opening the box at A, the observer there finds a right-hand glove (or a left-hand glove), then he or she can immediately state with certainty that at B a left-hand (or a right-hand) glove, respectively, has been or will be found. The observer at B can make the corresponding statement, if he or she has opened the box. *Correlation at a distance* is an everyday experience.

Of course, the type of glove was already determined in *reality* even before the boxes at A and B were opened. The gloves are *classically correlated*. Discovery of the type of glove at A had no influence either on the glove at A nor on the one at B (and correspondingly for opening of the box at B). The physical reality at the location of an observer is not changed by a distant experiment carried out by a partner. For gloves, there are only *local* occurrences. Although complete information about the other glove is obtained on opening one box, no transfer of information took place. The complete correlation is due simply to the fact that the 2-glove system was *prepared* initially as a pair of gloves. It is in the state "pair of gloves".

We have described a situation of classical physics in the case of the gloves, which – insofar as concerns the results of measurements – corresponds exactly to the situation of the pair of photons for parallel orientations of the analysers. The central question is therefore: can perhaps the 2-photon experiment also be described within a purely classical theory, i.e. without reference to quantum mechanics? The example of the gloves suggests already how we should search for differences between classical and quantum systems. For photons, one need not carry out the measurements at A and at B with the same analyser setup. The analysers can be rotated relative to one another, as we described in the previous section. With gloves, one can always ask only the same question at A and at B, "Left-hand or right-hand glove?". A question about the glove states $\frac{1}{\sqrt{2}}(|\text{left}\rangle \pm |\text{right}\rangle)$ is not possible. Such superpositions do not exist for classical systems (cf. Schrödinger's cat in Sect. 15.2.2). There is no way to carry out a "rotated" measurement. Certainly, pairs of gloves are too simple a model for photons. We will have to investigate much more complex classical models. In order to find a generally-valid answer to our question, we first want to examine in the example of the gloves which principles hold for a classical theory.

10.3 Local Realism

We refer conceptually to the article by A. Einstein, B. Podolsky und N. Rosen [EPR 35], and characterise classical physics by the two principles of *local realism* which we can observe from the pair of gloves:

Physical reality[1] Properties (e. g. energy) of physical systems are those physical quantities whose value can be predicted with certainty before carrying out the corresponding measurement (e. g. a measurement of the energy). These properties are in reality already determined

[1] Elements of reality: "If without in any way disturbing a system, we can predict with certainty the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity. This means that this physical quantity has a value independent of whether we measure it or not." [EPR 35]

before the measurement. The system "has them". They are *elements of physical reality*. Their values are independent of whether they are measured or not. We call this reality concept of classical physics *Einstein reality*.

Locality[2] Physical reality can be described in a *local* manner. This means that every system has its properties, independently of which interventions are carried out on other, spatially separated systems. There is no action at a distance. This concept of locality is called *Einstein locality*.

Hidden variables We assume – as did the authors of the EPR article also – that the experimental statements of the quantum theory are correct. If one furthermore assumes that Einstein reality and Einstein locality are appropriate descriptions of all physical systems, then all possible properties of each system are always present. The quantum theory affirms a different conclusion only because it is not structured finely enough to capture all of reality. If this assumption is justified, then quantum theory is not incorrect, but it is *incomplete*[3]. There are elements of reality which are not reflected in this theory (compare Sect. 2.2). They do not appear in it and are therefore *hidden variables* for the quantum theory.

As a *classical theory* we want to consider a theory which – in contrast to the quantum theory – fulfills the two requirements of local realism. This is true of theories which are usually termed classical, such as relativistic mechanics, electrodynamics, etc. We will attempt within this framework to introduce explicitly the previously hidden variables in order to construct a local-realistic and thus *classical alternative theory* to the quantum theory, which can give an account of all the experimentally-observed phenomena in the quantum regime. Briefly, we want to test the assertion: *There is only classical physics*. We pursue this programme and attempt to describe photons and spin-$1/2$ objects so to speak in the same way as we can describe pairs of gloves. We shall see that this is doomed to failure.

10.4 Hidden Variables, Bell Inequalities and Contradictions of Experiments

Stochastic local-realistic theory A proponent of local realism considers the experiment described in Sect. 10.1 and asserts that he can also derive the results. We want to test his assertion. A set of hidden variables summarised by the symbol λ represents the "elements of reality" which occur in connection with the polarisation of a now classically-described object. λ is a real variable with a certain range of values. The properties of a particular object are characterised by a certain set of variables λ. We formulate a *stochastic local-realistic theory*. Particles with values of the variables λ are produced by a source with the probability density $\rho(\lambda)$, for which

$$\int \rho(\lambda) d\lambda = 1, \qquad \rho(\lambda) \geq 0 \tag{10.16}$$

holds.

[2] Locality: "The real factual situation of system A is independent of what is done with system B, which is spatially separated from the former." [EPR 35].

[3] Completeness: "In a complete theory there is an element corresponding to each element of reality". [EPR 35]

For a classical object characterised by λ at A and for the angle of rotation δ_1 of the analyser at A, it is predetermined whether the direction of polarisation x^A or y^A will be observed (i.e. the corresponding detector will register a count). We attribute as in Sect. 10.2 a value $+1$ or -1 to the result of a measurement. Then there exists correspondingly an unambiguous function $S_A^\lambda(\delta_1)$ of λ and δ_1 with the values $+1$ or -1. Analogously, there is a possibly different function $S_B^\lambda(\delta_2)$, which for a given λ and a given rotation δ_2 at B determines unambiguously the measured value $+1$ or -1 there.

$$S_A^\lambda(\delta_1) = \left\{ \begin{array}{c} +1 \\ -1 \end{array} \right\}, \qquad S_B^\lambda(\delta_2) = \left\{ \begin{array}{c} +1 \\ -1 \end{array} \right\} . \tag{10.17}$$

The classical correlation coefficient is then

$$\epsilon^{cl}(\delta_1, \delta_2) = \int \rho(\lambda) S_A^\lambda(\delta_1) S_B^\lambda(\delta_2) d\lambda \, . \tag{10.18}$$

The fact that it is written as a product of the two functions $S_A^\lambda(\delta_1)$ and $S_B^\lambda(\delta_2)$ expresses locality. Through the formulation in terms of hidden variables λ, a local-realistic theory for the correlation coefficients has been created.

The Bell inequality In a first step, we adapt this up to now very generally formulated theory to the measured results of quantum mechanics which it must reproduce. In the special case of parallel orientation of the analysers, the well-confirmed complete correlation of the results

$$\epsilon^{cl}(\delta, \delta) = \int \rho(\lambda) S_A^\lambda(\delta) S_B^\lambda(\delta) d\lambda = 1 \tag{10.19}$$

must be valid. With Eqs. (10.16), (10.17), and (10.18), it follows that

$$S_A^\lambda(\delta) = S_B^\lambda(\delta) =: S^\lambda(\delta) \, . \tag{10.20}$$

The two observables have the same functional dependence on the hidden variables λ and on the angle δ.

We next go to a more general situation and consider three orientations, δ_1, δ_2 and δ_3. Measurements are carried out with the following orientations of the analysers at A and B: (δ_1, δ_2), (δ_2, δ_3) and (δ_1, δ_3). It is found to be expedient to write the following expression and to rearrange it using Eq. (10.17):

$$S^\lambda(\delta_1) S^\lambda(\delta_2) - S^\lambda(\delta_1) S^\lambda(\delta_3) = \underbrace{S^\lambda(\delta_1) S^\lambda(\delta_2)}_{=\pm 1} \underbrace{[1 - S^\lambda(\delta_2) S^\lambda(\delta_3)]}_{\geq 0} \, . \tag{10.21}$$

Integration leads with this expression and with Eqs. (10.16) and (10.17) to

$$\begin{aligned} & \left| \int \rho(\lambda) \{ S^\lambda(\delta_1) S^\lambda(\delta_2) - S^\lambda(\delta_1) S^\lambda(\delta_3) \} d\lambda \right| \\ & = \left| \int \rho(\lambda) S^\lambda(\delta_1) S^\lambda(\delta_2) [1 - S^\lambda(\delta_2) S^\lambda(\delta_3)] d\lambda \right| \\ & \leq \int \rho(\lambda) | \underbrace{[1 - S^\lambda(\delta_2) S^\lambda(\delta_3)]}_{\geq 0} | d\lambda = 1 - \int \rho(\lambda) S^\lambda(\delta_2) S^\lambda(\delta_3) d\lambda \, . \end{aligned} \tag{10.22}$$

Equation (10.22) implies for the classical correlation functions

$$|\epsilon^{cl}(\delta_1, \delta_2) - \epsilon^{cl}(\delta_1, \delta_3)| \leq 1 - \epsilon^{cl}(\delta_2, \delta_3) \, . \tag{10.23}$$

This is *the Bell inequality* ([Bel 64]).

Conflict with quantum theory What is the result predicted by quantum theory? We investigate spin-$1/2$ particles in the state $|\Phi_+^{AB}\rangle$ and choose the following orientations for the analysers: $\delta_1 = 60°, \delta_2 = 120°, \delta_3 = 180°$. Then we obtain for the correlation coefficients

$$\epsilon^{AB}(\delta_1, \delta_2) = \frac{1}{2}, \quad \epsilon^{AB}(\delta_1, \delta_3) = -\frac{1}{2} \quad \epsilon^{AB}(\delta_2, \delta_3) = \frac{1}{2}\,. \tag{10.24}$$

This expression cannot fulfill Eq. (10.23). Inserting it leads to

$$1 \leq \frac{1}{2}\,, \tag{10.25}$$

i. e. *The Bell inequality is not obeyed. Quantum theory and local-realistic theories lead to contradictory results.* For photons, one chooses an angle of $30°$ instead of $60°$ and likewise is led to a violation of the Bell inequality.

The CHSH inequality We discuss an additional combination of rotation angles. Reference to an experimental result, such as we built into Eq. (10.20), is not necessary here. At A, the measurements are carried out with the orientations α_1 and α_2, and at B with β_1 and β_2. We consider the following combination of the corresponding functions S and take note of the different values which can result owing to $S^\lambda_{A,B} = \pm 1$:

$$\begin{aligned} \{S_A^\lambda(\alpha_2)\,\underbrace{[S_B^\lambda(\beta_1) + S_B^\lambda(\beta_2)]}\, + S_A^\lambda(\alpha_1)\,\underbrace{[S_B^\lambda(\beta_1) - S_B^\lambda(\beta_2)]}\} \quad &=: \quad \{\ldots\} \\ \pm 2 \quad &\longleftrightarrow \quad 0 \\ 0 \quad &\longleftrightarrow \quad \pm 2 \end{aligned} \tag{10.26}$$

We thus have

$$|\{\ldots\}| = 2 \tag{10.27}$$

and therefore, with Eq. (10.16),

$$\left| \int \rho(\lambda)\{\ldots\} d\lambda \right| \leq \int \rho(\lambda) |\{\ldots\}| d\lambda = 2 \int \rho(\lambda) d\lambda = 2\,. \tag{10.28}$$

For the classical correlation coefficients ϵ^{cl} of the measurements with the various angles, this means that:

$$S^{cl} := |\epsilon^{cl}(\alpha_2, \beta_1) + \epsilon^{cl}(\alpha_2, \beta_2) + \epsilon^{cl}(\alpha_1, \beta_1) - \epsilon^{cl}(\alpha_1, \beta_2)| \leq 2\,. \tag{10.29}$$

This is the *CHSH inequality*, named for J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt [CHS 69].

All the inequalities for the correlation coefficients in stochastic theories with hidden parameters (stochastic local-realistic theories) are usually combined under the name *Bell inequalities*. *They are inequalities obtained from classical physics.* Their significance for quantum mechanics becomes clear only when one compares predictions based on the Bell inequalities (as in Eq. (10.25)) with those of quantum theory, and both with the actual experimental results.

The clash with quantum theory For the corresponding measurements on correlated photons in the entangled state $|\Phi_+^{AB}\rangle$, one chooses angles which differ by $22,5°$ (see Fig. 10.2b): $\alpha_1 = 22,5°, \beta_1 = 45°, \alpha_2 = 67,5°,$ and $\beta_2 = 90°$. This leads with Eq. (10.11) to:

$$\epsilon^{AB}(\alpha_1,\beta_1) = \epsilon^{AB}(\alpha_2,\beta_1) = \epsilon^{AB}(\alpha_2,\beta_2) = \cos\frac{\pi}{4} = \frac{1}{\sqrt{2}} \tag{10.30}$$

$$\epsilon^{AB}(\alpha_1,\beta_2) = \cos\frac{3\pi}{4} = -\frac{1}{\sqrt{2}}\,. \tag{10.31}$$

(For spin-$\frac{1}{2}$ particles, the angles chosen are twice as large.) The quantum-mechanical computation of the correlation coefficients as in Eq. (10.29) yields for the entangled state

$$S^{e.s.} = 2\sqrt{2}\,. \tag{10.32}$$

Comparison with the CHSH inequality leads to the conclusion: *Quantum mechanics violates the CHSH inequality*.

Experimentum crucis[4] We have described two experimental setups for which the quantum theory on the one hand and theories with hidden variables on the other lead to contradictory predictions. In such a situation, experimental results can allow a choice to be made. The experiments verify the predictions of quantum mechanics (compare Sect. 10.8). *All local-realistic alternative theories to quantum mechanics are thus refuted.* This is a result with far-reaching consequences. The underlying reason is to be found in the fact that the EPR correlations cannot be simulated in the entangled Bell state by introducing hidden local-realistically interpreted variables. They are genuine quantum correlations. For a non-local alternative theory to the quantum theory, see Sect. 10.8.

10.5 Separable Mixtures Obey the Bell Inequality

Separable mixtures such as those as we introduced in Sect. 8.1 can be simulated by mixtures which have been locally prepared via LOCC. As quantum mixtures, they do not obey the reality condition. Nevertheless, they are referred to as classically-correlated mixtures. We want to show that this is justified, since they obey the CHSH inequality. We consider to this end separable mixtures of bipartite states

$$\rho^{AB} = \sum_i p_i \rho_i^A \otimes \rho_i^B, \qquad 0 < p_i \leq 1, \qquad \sum_i p_i = 1 \tag{10.33}$$

and again compute the quantum-mechanical correlation coefficients (compare Eq. (10.9))

$$\begin{aligned} \epsilon^{AB}(\alpha,\beta) &= tr_{AB}[\rho^{AB} E^A(\alpha) \otimes E^B(\beta)] \\ &= \sum_i p_i A_i(\alpha) B_i(\beta) \end{aligned} \tag{10.34}$$

[4] An experiment which permits one theory among several to be verified and the others to be refuted is called an *experimentum crucis* (experiment of the cross). This term is due to Bacon (1561-1626). The addendum 'cross' is intended to recall a fork in the road (or a crossroads), which in Bacon's time was often marked by a cross.

with

$$A_i(\alpha) := tr_A[\rho_i^A E^A(\alpha)], \qquad B_i(\beta) := tr_B[\rho_i^B E^B(\beta)]\,. \tag{10.35}$$

The magnitudes of the expectation values of operators with eigenvalues equal to ± 1 cannot exceed 1; therefore, we have:

$$|A_i(\alpha)| \leq 1, \qquad |B_i(\beta)| \leq 1\,. \tag{10.36}$$

The evaluation of the correlation coefficient leads to

$$\begin{aligned} S^{s.m.} &= \left|\epsilon^{AB}(\alpha_2,\beta_1) + \epsilon^{AB}(\alpha_2,\beta_2) + \epsilon^{AB}(\alpha_1,\beta_1) - \epsilon^{AB}(\alpha_1,\beta_2)\right| \qquad (10.37)\\ &= \left|\sum_i p_i\{A_i(\alpha_2)B_i(\beta_1) + A_i(\alpha_2)B_i(\beta_2) + A_i(\alpha_1)B_i(\beta_1) - A_i(\alpha_1)B_i(\beta_2)\}\right| \\ &\leq \sum_i p_i\{|A_i(\alpha_2)B_i(\beta_1) + A_i(\alpha_2)B_i(\beta_2)| + |A_i(\alpha_1)B_i(\beta_1) - A_i(\alpha_1)B_i(\beta_2)|\}\,. \end{aligned}$$

As a result of the separability, only products occur here. We estimate the value of the expression in curved brackets $\{\ldots\}$ in the same way as in Eq. (10.36):

$$\{\ldots\} \leq |B_i(\beta_1) + B_i(\beta_2)| + |B_i(\beta_1) - B_i(\beta_2)| \leq 2 \tag{10.38}$$

and obtain with Eq. (10.33) for the separable mixture

$$S^{s.m.} \leq 2\,. \tag{10.39}$$

Comparison with the CHSH inequality (10.29) shows *that for bipartite systems, every separable mixture of quantum states (and thus also every pure product state) obeys the CHSH inequality*. Entanglement is therefore a necessary condition for the violation of the Bell inequality. It is, however, not a sufficient condition (see Sect. 10.8).

10.6 Entanglement Witnesses*

We carried out the computations in the previous section in analogy to those in the sections preceding it. In order to formulate the result once more in a different way, we introduce vectors $\boldsymbol{a_1}$ and $\boldsymbol{a_2}$ (or $\boldsymbol{b_1}$, $\boldsymbol{b_2}$) to denote the directions α_1 and α_2 (or β_1, β_2). Making use of the *Bell-CHSH observable*

$$B^{AB} := \boldsymbol{\sigma}^A \boldsymbol{a_2} \otimes \boldsymbol{\sigma}^B(\boldsymbol{b_1} + \boldsymbol{b_2}) + \boldsymbol{\sigma}^A \boldsymbol{a_1} \otimes \boldsymbol{\sigma}^B(\boldsymbol{b_1} - \boldsymbol{b_2})\,, \tag{10.40}$$

we can then summarise equations (10.37) through (10.39) in the form

$$|\mathrm{tr}_{AB}[\rho^{AB} B^{AB}]| \leq 2\,. \tag{10.41}$$

Here, we have made use of Eqs. (10.34) and (10.14). All separable mixtures obey Eq. (10.41).

*The sections marked with an asterisk * can be skipped over in a first reading.

We reformulate this result: the mean value of the observable $\hat{W}^{AB} := 2\mathbb{1}^{AB} + B^{AB}$ formed with the help of B^{AB} fulfills for *all* separable mixtures ρ_{sep}^{AB} the relation

$$\mathrm{tr}_{AB}[\rho_{sep}^{AB}\hat{W}^{AB}] \geq 0\,. \tag{10.42}$$

For the entangled state $|\Phi_+^{AB}\rangle$, Eq. (10.32) means that

$$\mathrm{tr}_{AB}[|\Phi_+^{AB}\rangle\langle\Phi_+^{AB}|\hat{W}^{AB}] < 0\,. \tag{10.43}$$

Therefore, if one carries out the Bell inequality experiments on a system in the state ρ^{AB} and the results show that the inequality (10.42) is not obeyed, then the state is entangled.

This observation can be generalised. It suggests that we establish a universal concept. An *entanglement witness* is an observable W^{AB} which reveals the presence of entanglement in mixtures $\rho^{AB} \in \mathcal{H}_m^A \otimes \mathcal{H}_n^B$. W^{AB} fulfills the conditions

$$\mathrm{tr}_{AB}[\rho^{AB}W^{AB}] \quad \begin{cases} \geq 0 & \text{for all non-entangled } \rho^{AB} \\ < 0 & \text{for at least one } \rho^{AB}\,. \end{cases} \tag{10.44}$$

It is not excluded that the first inequality is obeyed also by one or more entangled states. Thus, the result of the measurement of W^{AB} on a system in the state ρ^{AB} enables us to draw the following conclusions:

$$\mathrm{tr}_{AB}[\rho^{AB}W^{AB}] \quad \begin{cases} < 0 & \Rightarrow \rho^{AB} \text{ is entangled} \\ \geq 0 & \Rightarrow \text{no conclusion about the entanglement} \\ & \quad \text{of } \rho^{AB} \text{ is possible.} \end{cases} \tag{10.45}$$

The special observable introduced above, $\hat{W}^{AB}$ on $\mathcal{H}_1^A \otimes \mathcal{H}_2^B$, is an example of an entanglement witness.

If one finds an entanglement witness W^{AB} for a state ρ^{AB}, such that Eq. (10.45) is fulfilled, then the state is entangled. But does an entanglement witness exist for every entangled state? This is guaranteed by the following theorem, which we cite without proof [HHH 96]: *A state ρ^{AB} is entangled if and only if there is a Hermitian operator W^{AB}, for which all separable states obey the inequality $tr_{AB}[\rho_{sep}^{AB}W^{AB}] \geq 0$ and for which ρ^{AB} fulfills the relations $tr_{AB}[\rho^{AB}W^{AB}] < 0$.* No limitations on $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$ were made in this theorem.

The theorem can have practical utility in finding out whether a state is entangled or not, as shown by the special case $\hat{W}^{AB}$ discussed initially. In Sect. 10.8, we give a further example which is based on a Bell measurement. An entanglement witness is a non-local observable. At first glance, it would therefore appear difficult to establish entanglement in this way experimentally. In fact, this is not the case, since the statements made by Eq. (10.45) are based on mean values. As we saw in Sect. 9.2.2, they can be determined merely by local projection measurements on the subsystems. In practice, one tests with the aid of entanglement witnesses whether a preparation procedure has in fact produced a particular entangled state. The entanglement witness is constructed specifically for this state. The observable $\hat{W}^{AB}$ in Eq. (10.43) is an example of this. The introduction of entanglement witnesses does not, however, solve the separability problem, since we do not know all the possible entanglement witnesses.

10.7 3-Particle Entanglement and Quantum Locality

10.7.1 The GHZ State

D. M. Greenberger, M. A. Horne and A. Zeilinger (GHZ) have shown in a non-statistical manner, which is independent of the Bell approach, that the quantum theory and local realism are not mutually compatible ([GHZ 89], [GHZ 90]). We give their arguments here with reference to the spins of three quantum objects which are located at three different places, A, B, and C. They are mutually entangled. Their composite state is in $\mathcal{H}_2^A \otimes \mathcal{H}_2^B \otimes \mathcal{H}_2^C$ and is supposed to be given by

$$|\psi^{ABC}\rangle = \frac{1}{\sqrt{2}}(|0^A, 0^B, 0^C\rangle + |1^A, 1^B, 1^C\rangle) \; . \tag{10.46}$$

$|0\rangle$ and $|1\rangle$ are again the eigenvectors of σ_z with the eigenvalues $+1$ and -1. The state (10.46) of the tripartite system is called the *GHZ state*. It is symmetric with respect to exchange of the tags A, B and C.

10.7.2 Local Realism and Quantum Theory at Odds

We first determine again the result of local quantum-mechanical spin measurements on the three subsystems. For this, different polarisation directions are chosen. For example, in the yyx measurement, the observables σ_y^A, σ_y^B and σ_x^C are each measured locally on a single composite system. The order in which the measurements are performed is unimportant. To find the result of the measurements, it is expedient to introduce the eigenvectors of each observable as a basis in the spaces $\mathcal{H}_2^A, \mathcal{H}_2^B$ and $\mathcal{H}_2^C$ and to expand the state vector $|\psi^{ABC}\rangle$ in this basis. Then, after an intermediate calculation, we find

$$|\psi^{ABC}\rangle = \frac{1}{2}(|0_y^A, 1_y^B, 0_x^C\rangle + |1_y^A, 0_y^B, 0_x^C\rangle + |0_y^A, 0_y^B, 1_x^C\rangle + |1_y^A, 1_y^B, 1_x^C\rangle) \; . \tag{10.47}$$

The measured values are denoted by s_x and s_y. They are always $+1$ (eigenvector $|0\rangle$) or -1 (eigenvector $|1\rangle$). As can be read off from Eq. (10.47), the possible combinations of the correlated measured values for each individual tripartite system obey the relation

$$s_y^A s_y^B s_x^C = -1 \; . \tag{10.48}$$

The symmetry under exchange of the tags A, B and C leads to

$$s_x^A s_y^B s_y^C = -1, \qquad\qquad s_y^A s_x^B s_y^C = -1 \; . \tag{10.49}$$

These are the results of quantum theory, which are confirmed by experiments.

The situation is again the same as in the previous sections. An exponent of local realism observes the results of the measurements and claims that he can explain them completely within the framework of his approach. Is he right? We need to verify whether the experimental results by themselves could not also be accounted for by a classical model. To this end, we consider a classical overall system which is composed of three subsystems S^A, S^B

and S^C. We can carry out x and y measurements on all three systems. They are completely characterised by the requirements which result from the experiments: the possible measured values $s_{x,y}^{A,B,C}$ are always $+1$ or -1 and the conditions (10.48) and (10.49) must be fulfilled for the products. The following property of the system can then be read off: the result of the x measurement on one of the subsystems can be predicted with certainty when the results of the y measurements on the other two subsystems are known. In order to determine e.g. the result s_x^A of the x measurement on S^A, one need only carry out a y measurement on S^B and S^C. Analogously, the result of a y measurement can be predicted if one carries out an x measurement or a y measurement on the other two subsystems.

Since we wish to interpret this observation from the point of view of a local realistic theory, there is a consequence: the outputs of all the x and y measurements are predetermined in our classical model, e.g. via hidden variables. For a single tripartite system, the individual values $s_x^{A,B,C}$ are already fixed before the measurement. The values $s_{x,y}^{A,B,C}$ are therefore the same in the different Eqs. (10.48) and (10.49). We multiply the left sides of the three equations and obtain with $s_y^A s_y^A = s_y^B s_y^B = s_y^C s_y^C = 1$

$$s_x^A s_x^B s_x^C \stackrel{cl}{=} -1\,. \tag{10.50}$$

In every local realistic theory, this is the prediction for an xxx measurement.

We still have to compute the corresponding prediction of quantum theory. For a measurement of the observables σ_x^A, σ_x^B and σ_x^C, we decompose $|\psi^{ABC}\rangle$ in terms of the eigenvectors of σ_x

$$|\psi^{ABC}\rangle = \frac{1}{2}(|0_x, 0_x, 0_x\rangle + |0_x, 1_x, 1_x\rangle + |1_x, 0_x, 1_x\rangle + |1_x, 1_x, 0_x\rangle) \tag{10.51}$$

and find

$$s_x^A s_x^B s_x^C \stackrel{GHZ}{=} +1\,. \tag{10.52}$$

Thus, there is a clear-cut contradiction between the predictions of local-realistic theories and of quantum theory. If experiments confirm the quantum-mechanical equation (10.52) – which is indeed the case (see Sect. 10.8) – then local realism is refuted.

Why can one not derive a relation (10.52) in a similar manner in quantum mechanics from Eqs. (10.48) and (10.49)? Equation (10.47) shows that the quantum-mechanical measured results are correlated. When local measurements on a GHZ system yield e. g. $s_y^B = +1$ and $s_x^C = -1$, then one finds $s_y^A = +1$. For $s_y^B = +1$ and $s_x^C = +1$, one finds with an identically-prepared system $s_y^A = -1$, etc. The result s_y^A is in general not predetermined for a tripartite system in the GHZ state. This may differ from the relations (10.48) and (10.49), which refer to different measurements. The same holds for s_y^B and s_y^C.

We have shown the following: *Three objects are prepared in such a way that in the framework of quantum theory, the resulting state is the GHZ state (10.46). If the quantum-mechanical predictions for the local polarisation measurements as described above prove correct, then every local-realistic theory for these systems is refuted.* Quantum theory can therefore not be replaced by such theories.

We have seen in Sect. 3.6 that the two linear polarisations and the circular polarisation can be formulated analogously to the spin in $\mathcal{H}_2$. For photonic GHZ states, the contradiction can

be demonstrated in a similar manner. *For photons, the experimental results in fact confirm the quantum-mechanical predictions and thus refute the local-realistic approach.* A summary of the experiments can be found e. g. in [PZ 02].

Finally, we point out some differences from the Bell arguments. The CHSH inequality makes statements about classical expectation values and is a direct result of local realism. It makes a statement about classical physics without referring to the quantum theory. The GHZ argument, in contrast, is based (like the conjuring trick in Sect. 7.7) on the failure of the attempt to simulate quantum-mechanical results (10.48) and (10.49) with a local-realistic method. The contradiction with quantum theory is not probabilistic, but rather direct. In both cases, experiments verify the quantum theory.

10.8 Complementary Topics and Further Reading

- The classics: [EPR 35], [Bel 64], [CHS 69], [Boh 51], [GHZ 89], [GHZ 90].
- Review articles: [HS 91], [Per 93, Chap. 6], [Hom 97, Chap. 4], [AS 99], [Aul 00, Chap. IX], [WW 01].
- On Einstein's criticism of quantum theory: [Hom 97, Chap.8].
- An introduction to the debates between N. Bohr and A. Einstein on the basis of quantum theory: [Hel 06]. The correspondence between A. Einstein and M. Born is also worth reading: [EB 69].
- There are entangled states that do not violate any of the known inequalities which are of the type of the Bell inequalities: [WW 00]. This is interesting in view of the entanglement witnesses $\hat{W}^{AB}$ discussed in Sect. 10.6.
- Books with review articles on the Bell inequalities and relevant experiments: [BZ 02], [Asp 02].
- An experiment on the violation of the Bell inequality using two ions is described in [RKM 01]. There, one also finds numerous literature references to experiments with photons.
- Objections against the experimental demonstrations of the violation of the Bell inequality have been raised on the grounds that they contain *loopholes* which would permit local-realistic interpretations. The history of this problem is given in [Asp 99].
- Discussions of entanglement witnesses in a wider context: [LBC 00], [HHH 01], [Ter 02], [Cir 02].
- The *Bohm theory*, which is also called the *de Broglie-Bohm theory* or *Bohmian mechanics*, is an alternative theory to conventional quantum mechanics (cf. [Boh 83]). It is thus not just another interpretation, but rather a theory which is different from quantum theory. It is based on the introduction of non-local hidden variables and can reproduce all the results of nonrelativistic quantum mechanics, in particular also the violation of the Bell

inequality. The wavefunction is a physically real object (comparable to for example the gravitational field). It guides the particles along their orbits. All its laws are completely deterministic. The differential equation for the guide waves is the Schrödinger equation. Attempts at a relativistic and quantum-field generalisation of the Bohm theory have met with serious problems. The fact that nearly all physicists prefer standard quantum theory to the Bohm theory, in spite of its agreement with experimental results, is due to "meta-theoretical" arguments in the evaluation of the theories (see Sect.2.4), and especially also to pragmatic arguments (poorly-developed and clumsy formalism). In [Pas 05], a brief description of the Bohm theory and of the objections raised against it are given, together with a rather complete literature list.

10.9 Problems for Chapter 10

Prob. 10.1 [for 10.1]: Give an alternate (direct) proof of Eq. (10.15).

Prob. 10.2 [for 10.4]: Prove Eq. (10.20).

Prob. 10.3 [for 10.4]: Measurements are carried out on linearly-polarised photons in the state $|\varphi^A\rangle$ with an analyser A oriented in the directions x^A and y^A. Formulate a model for this situation with hidden variables, i. e. give expressions for λ, $\rho(\lambda)$ and S_A^λ which reproduce the measured results.

Prob. 10.4 [for 10.7]: Design a quantum circuit which transforms the state $|0^A, 0^B, \ldots, 0^J\rangle$ in $\mathcal{H}_2^A \otimes \mathcal{H}_2^B \otimes \ldots \otimes \mathcal{H}_2^J$ into the GHZ state

$$|GHZ\rangle = \frac{1}{\sqrt{2}}(|0^A, 0^B, \ldots, 0^J\rangle + |1^A, 1^B, \ldots, 1^J\rangle) \; .$$

Prob. 10.5 [for 10.6]: Show that for mixtures ρ^{AB} in $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$, the operator $W^{AB} = \mathbb{1}^{AB} - 2|\Phi_+^{AB}\rangle\langle\Phi_+^{AB}|$ is an entanglement witness.

© 2007 WILEY-VCH Verlag GmbH & Co.

11 Working with Entanglement

Entanglement forms the basis for new effects and their technical applications. We discuss some examples. In the second part of this section, we will manipulate entanglement. The resource "entanglement" cannot be implemented at its maximum strength in practice. It is therefore important to have access to a method of amplifying entanglement. In order to control this method qualitatively, we require measures of entanglement.

11.1 Quantum Cryptography

11.1.1 The Vernam Coding

Alice would like to send a coded message to Bob, which is to be kept perfectly secret. No one besides Bob should be able to decode it. In the method suggested by Vernam in 1926 [Ver 26], a *source text*, which is supposed to be available in digital form as a series of (0,1) bits of length n, is coded using a *key*. This produces the *cryptogram*. The key itself consists of a random (0,1) series which likewise has the length n. For coding, the numbers from the source text and the key are added modulo 2 term by term. We give an example:

Source text	01101100
Key	10000110
Cryptogram	11101010

The cryptogram is then sent to Bob, who is presumed to be already in possession of the key. The decoding process consists of Bob's adding the key to the coded message, modulo 2. Due to $x + 0 + 0 = x + 0 = x$ and $x + 1 + 1 = x + 0 = x$, the source text is again obtained:

Cryptogram	11101010
Key	10000110
Source text	01101100

Since the key contains a random sequence, the coded text becomes completely independent of the source text. It can be transmitted openly by Alice to Bob. For an eavesdropper who does not have the key, the cryptogram contains no usable information. The decisive point is that the key be known only to Alice and Bob, that it is indeed a genuinely random sequence of the same length as the source text, and in particular, that it be used only once (one-time pad system). Under these conditions, the Vernam code cannot be broken [Sha 49].

The problem with this procedure is that Bob and Alice must exchange a new key for every message they want to send, and that in this process, it must be guaranteed that the key cannot

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

fall into the hands of an eavesdropper. We want to show that it is possible to accomplish such a key transmission using quantum systems. There are a whole series of quantum-cryptographic methods (compare Sect. 11.8). Here, we treat only two types of methods. One of these makes use of the special properties of the quantum-mechanical measurement process, and the other employs the non-local EPR correlations. The series of instructions for carrying out a particular cryptographic scheme is called a *protocol*.

11.1.2 The B92 Protocol

The B92 protocol, named for the work of C. H. Bennett [Ben 92] from the year 1992, makes use of two non-orthogonal quantum states for the transmission of the key and utilises the following typical properties of individual quantum systems: (i) There is no measurement procedure which can distinguish between the two states. (ii) The states cannot be cloned by a quantum copier. (iii) A measurement in general modifies a quantum state.

The process of key transmission We want to use photons to implement the B92 protocol. Alice has two filters which can linearly polarise the photons vertically in the state $|V\rangle$, or at an angle of $-45°$ to the vertical, i. e. in the state $|V'\rangle$ (see Sect. 3.6). The relation $|\langle V|V'\rangle|^2 = \frac{1}{2}$ holds. Alice uses a long random binary series, e. g. $1, 0, 0, 1, 1, 0, \ldots$, and every time a 0 occurs, she generates a photon of polarisation $|V\rangle$, while for each 1, she chooses the polarisation $|V'\rangle$. The photons propagate without disturbance to Bob's location.

Bob has a measuring apparatus in the form of a detector behind a polarisation filter. He can choose one of two orientations for the analyser filter: a horizontal polarisation $|H\rangle$, corresponding to the projection operator $P_1 = \mathbb{1} - |V\rangle\langle V| = |H\rangle\langle H|$, and a polarisation rotated by $-45°$ with respect to the horizontal, corresponding to the projection operator $P_0 = \mathbb{1} - |V'\rangle\langle V'| = |H'\rangle\langle H'|$. Bob also has a long binary sequence at his disposal, which is independent of the sequence used by Alice, e. g. $0, 0, 1, 0, 1, 0, \ldots$. The numbers 0 and 1 must occur with equal frequencies in both sequences. The i th number of Bob's sequence determines the type of measurement that he will make on the i th photon prepared by Alice. If the number 0 occurs in Bob's sequence, he places the analyser direction along $|H'\rangle$ (corresponding to P_0), and in the case of a 1, he places it along $|H\rangle$ (corresponding to P_1). In each case, he notes whether or not his detector registers a count.

Owing to the orthogonality of the vectors, we have $P_0|V'\rangle = 0$ and $P_1|V\rangle = 0$. When Alice and Bob find different numbers in their respective sequences, Bob's detector thus does not register a count. If Alice and Bob find the same numbers, then the detector registers a count with the probability $\frac{1}{2}$ due to $\langle V|P_0|V\rangle = \langle V'|P_1|V'\rangle = \frac{1}{2}$, thus in 50% of the cases. An example is shown in Table 11.1. "yes/no" means that in this arrangement the results 'yes' or 'no' are each possible with a probability $\frac{1}{2}$. The case which is in fact observed is underlined.

In the next step, Bob informs Alice via a public channel which photons produced a count in his detector. He does not inform her of the corresponding analyser position. In the example shown in the table, these are the photons numbered 2 and 6. Then Alice and Bob have the same sequence of numbers, $0, 0, \ldots$, which is known only to the two of them and which they can use as a key. As desired, it is a random sequence of the numbers 0 and 1, but it consists on the average of only 25% of the numbers from the original sequences.

Table 11.1: B92 Protocol for Quantum Cryptography.

Photon number	1	2	3	4	5	6
Alice's sequence	1	0	0	1	1	0
Polarisation generated	$\|V'\rangle$	$\|V\rangle$	$\|V\rangle$	$\|V'\rangle$	$\|V'\rangle$	$\|V\rangle$
Bob's sequence	0	0	1	0	1	0
Analyser position	$\|H'\rangle$	$\|H'\rangle$	$\|H\rangle$	$\|H'\rangle$	$\|H\rangle$	$\|H'\rangle$
Detector count	no	yes/no	no	no	yes/no	yes/no

Defence against eavesdropping We want to discuss briefly the question of security in this process of generating a common key. Let us assume that a third person named *Eve*[1] attempts by *eavesdropping*, i. e. by intercepting the photons and making her own polarisation measurements, to obtain the key; she then has the problems listed above under (i) – (iii). She cannot determine the polarisation state of a single photon in one measurement with certainty. She cannot copy the state onto many photons, let one of them continue on, and then carry out measurements on the copies. Finally, she will modify the state of each photon by her measurement before it is sent on to Bob. This is the circumstance which makes it possible for Alice and Bob to determine whether a spy was at work. The changes in the state of the photons will cause Bob now and again to register a count in his analyser-detector system, even though the same numbers were not found at his location and at Alice's. Both Alice and Bob can become aware of this by exchanging a randomly-chosen portion of their key publicly and comparing. If eavesdropping is found to have occurred, then the entire key is rejected and the process of generating a key is begun again. A test for eavesdropping can additionally be made by verifying whether 25% of the photons contribute to the key.

Improving security In fact, the transmission of the quantum objects from Alice to Bob is susceptible to disturbances, even when no one is eavesdropping. The quantum channel is generally noisy. In order to generate a key in practice, Alice and Bob have to tolerate a certain degree of errors, without knowing whether they are not after all due to the influence of Eve. This situation occurs in all the quantum-mechanical cryptography methods. In the final step of the corresponding protocol, classical algorithms are therefore employed in order to correct the errors. In an error correction protocol, an attempt is made to obtain a shorter key. This is accompanied by a procedure which reduces Eve's information to a minimum (private amplification algorithm). Review articles on this topic are listed in Sect. 11.8.

Other 1-qubit protocols The B92 protocol, which is based on two non-orthogonal states, is susceptible to POVM measurements. These non-projective measurements will be introduced in Chap. 13. It has therefore become standard practice to use four states, as in the BB84 protocol. Another approach is the following: the symmetry of the Bloch sphere makes it attractive to employ the eigenstates of σ_x, σ_y and σ_z. The corresponding 6-state protocol allows a simplified security analysis (see Sect. 11.8).

[1] For "eavesdropper"

11.1.3 EPR Protocols

Quantum cryptography using Bell's theorem The experiment described in Sect. 10.4 with pairs of photons for testing the CHSH inequality can also be used for transmitting keys. The protocol contains the following steps, which are carried out either publicly or else in secret by Alice (at A) or by Bob (at B).

Public

At A and B, the four measurement orientations are predetermined: $\alpha_1 = 22,5°, \beta_1 = 45°$, $\alpha_2 = 67,5°, \beta_2 = 90°$ (as in Sect. 10.4, see Fig. 10.1)

At A and at B in secret

The source generates pairs of entangled photons in the Bell state $|\Phi_+^{AB}\rangle$. At A and B, independently of each other in completely random sequences, the analysers are oriented in each of the 4 directions. At A, as described in Sect. 10.1, the polarisation states $|x^A\rangle$ (measured value $+1$) or $|y^A\rangle$ (measured value -1) are measured, and correspondingly at B, $|x^B\rangle$ (measured value $+1$) or $|y^B\rangle$ (measured value -1).

The information as to which polarisation directions were chosen at A and at B for the individual pairs of photons is exchanged publicly.

At A and B, the measurement results are sorted according to whether the same or different analyser directions were used.

Those measurement results from *different* orientations are exchanged publicly, together with the numbers of the pairs of photons. Thereby, each partner can verify whether the quantum-mechanical value for the correlations from Eq. (10.32), $S^{e.s.} = 2\sqrt{2}$, was obtained. If not, the transmission series is abandoned, since it has been tapped. If the value $S^{e.s.}$ is obtained,

Then the results for the *same orientations*, which were not exchanged, are perfectly correlated (compare Sect. 10.1). They represent a completely random sequence of the numbers 0 and 1, which can be used by Alice and Bob as a key.

For key generation with spin-$\frac{1}{2}$ particles, orientations with twice the angles are chosen.

The individual subsystems at Alice's and Bob's locations are in the maximally-mixed states $\rho^A = \rho^B = \frac{1}{2}\mathbb{1}$. No information at all can be read out of them. As we have already seen in Chap. 9, the information is contained only in the correlations. If an eavesdropper carries out a measurement, e. g. on the objects being transmitted to Alice, then he transforms this subsystem into a pure state. We showed in Sect. 8.3.4 that this breaks the entanglement and produces a separable state. According to Sect. 10.5, separable states obey the CHSH inequality, i. e. the relation $|S| \leq 2$ holds. The explicit calculation leads in fact to $|S| \leq \sqrt{2}$ (see Problem 11.3). This value differs strongly from the quantum-mechanical result, $S^{e.s.} = 2\sqrt{2}$.

The BBM92 protocol[2] There is a very simple protocol which is based on EPR correlations and requires no reference to a Bell inequality. We again use pairs of photons in the rotationally-symmetric Bell state $|\Phi_+^{AB}\rangle$, which we can write in terms of the linear polarisations $|H\rangle, |V\rangle$ and $|H'\rangle, |V'\rangle$ from Sect. 3.6

$$|\Phi_+^{AB}\rangle = \frac{1}{\sqrt{2}}(|H^A, H^B\rangle + |V^A, V^B\rangle) = \frac{1}{\sqrt{2}}(|H'^A, H'^B\rangle + |V'^A, V'^B\rangle) \,. \quad (11.1)$$

Alice and Bob make their measurements independently of one another, selecting in a completely random manner either the polarisations $|H\rangle$ and $|V\rangle$ or the polarisations $|H'\rangle$ and $|V'\rangle$, which are rotated by -45° relative to these.

After a series of measurements on pairs of photons, Alice and Bob exchange information about the polarisation directions they have chosen for the individual pairs. The results belonging to different directions, and those in which a photon was lost, are eliminated. The remaining measurement results must be perfectly correlated, presuming that no eavesdropping has occurred. In order to test this, Alice and Bob again compare a sufficiently large subset of these measurements via a public channel. In the case that their comparison is positive, the sequence of the remaining results, which is the same for both, can be used as the key for coding.

11.1.4 The Scheme of Quantum Cryptography

The basic idea consists of permitting Bob and Alice to obtain the same key by making use of quantum systems as carriers. The message itself is transmitted via a publicly accessible channel after being coded. The protocol must be designed in such a way that Alice and Bob can each determine whether eavesdropping took place during transmission of the key. To this end, they make use of the fact that a quantum measurement by an eavesdropper would change the state of the quantum objects being transmitted, if it was not precisely an eigenstate of the observables being measured.

If the preparations of Alice and the results of the corresponding measurements by Bob agree with the theoretical prediction, no spying measurements have occurred and no information has been obtained by an eavesdropper. Alice and Bob exchange a portion of their results publicly and verify the agreement. If there is disagreement, the key transmission is rejected

[2] Named for the article [BBM 92].

and the protocol is repeated from the beginning. Otherwise, Alice and Bob have obtained a secret key which furthermore represents a perfectly random distribution, since it is based on quantum processes.

11.2 One Qubit Transmits Two Bits (Dense Coding)

In Chap. 6, we have seen that one can code a single bit in a qubit and can read it out again. Is it possible to transmit more classical information with a single qubit? We wish to show that *dense quantum coding* makes it possible to transmit two bits via one qubit. The trick here is that before the transmission, an entangled 2-qubit system was already established at Alice's and Bob's locations. For example, Alice has sent a qubit system (e. g. one photon from a pair) to Bob and kept its entangled partner. The state of the composite system can be e. g. $|\Phi_+^{AB}\rangle$. Bob thus obtains no information in this process. Furthermore, Alice and Bob have previously agreed on how they will associate two bits with the four Bell states $|\Phi_+^{AB}\rangle$, $|\Phi_-^{AB}\rangle$, $|\Psi_+^{AB}\rangle$, and $|\Psi_-^{AB}\rangle$.

In Sect. 9.2.1, we pointed out that, using the σ operators, one can transform a Bell state locally in a unitary fashion into any other Bell state. Alice is supposed to carry out the transformations $\mathbb{1}^A$ (trivial), σ_1^A, $i\sigma_2^A$, and σ_3^A on her qubit. In order to transmit the two bits of information, she thus transforms the entangled state $|\Phi_+^{AB}\rangle$ into the corresponding Bell state and sends her qubit system to Bob. Bob can thus make use of both subsystems and determine by means of a Bell measurement which Bell state is present. By the transmission of one qubit, two bits have been transferred to him.

Dense coding is however hard to implement. If the entangled pure state used is not maximally entangled, the amount of information transmitted decreases and approaches in the limiting case one bit. An important property of the procedure is its security with respect to eavesdropping. In the most unfavourable case, one bit can be read off the qubit system which Alice sends to Bob. Dense coding demonstrates once again the significance of entanglement as a resource for information transmission.

11.3 Quantum Teleportation

Alice is in possession of a classical object which is however unknown to her, e. g. an oddly-shaped iron ball which is locked up in a box. Bob would like to have a similar object. To meet his wish, Alice has to open the box and carry out optical measurements on the ball. She then transmits the information to Bob via a classical channel and he can manufacture a ball in the same state from a block of iron. If this is meant to succeed not only approximately, but rather precisely, then in principle infinitely many bits of information will have to be transmitted.

In the analogous quantum-mechanical situation, one can proceed in a similar manner. To determine the unknown state of a quantum object requires in principle an infinite number of measurements. If Alice knows the quantum state, she must still transmit infinitely many bits to Bob. This can already be seen in the example of qubit systems. Let us suppose that the qubit system which Alice has is in an eigenstate of $\boldsymbol{\sigma r}$. In order to describe the vector $\mathbf{r}$, one

requires a binary number with infinitely many digits. Only when he has complete knowledge of $\mathbf{r}$ can Bob repeat the preparation procedure exactly.

The problem to be solved is the same in both cases. Alice and Bob are supposed to install cooperatively a preparation procedure which carries out the following task: Alice has an object in a state which is unknown to Alice and to Bob. The procedure allows the preparation of an object at Bob's location which is in precisely the same state. For the procedures described so far, this requires the transmission of a large number of bits over a classical channel. This is unavoidable in the case of the classical state. For the transmission of the quantum state, one can however proceed in a much more subtle manner. Keeping in mind dense coding, it makes sense to again make use of an *information transmission assisted by entanglement*. In fact, the following protocol of *quantum teleportation* leads to the desired result (cf. Fig. 11.1):

Alice and Bob again share the Bell state $|\Phi_+^{AB}\rangle$. The subsystems are the quantum systems S^A and S^B, which are at Alice's and Bob's locations, respectively. Alice has an additional quantum system S^C in a pure state

$$|\varphi^C\rangle = a|0^C\rangle + b|1^C\rangle \tag{11.2}$$

with $|a|^2 + |b|^2 = 1$, which is unknown to Alice. This state $|\varphi\rangle$, not the quantum system S^C itself, is to be teleported to Bob. That means that we are seeking a procedure by means of which Bob's subsystem S^B can be prepared in the pure state $|\varphi^B\rangle$. S^B is then necessarily no longer entangled with any other system (cf. Sect. 8.3.4).

All together, there is a tripartite system. We carry our some algebraic manipulations on its state in $\mathcal{H}_2^C \otimes \mathcal{H}_2^A \otimes \mathcal{H}_2^B$. For these, we make use of the definition of the Bell states and of the properties of the Pauli operators, and in an intermediate step, we introduce a Bell basis in $\mathcal{H}_2^C \otimes \mathcal{H}_2^A$.

$$\begin{aligned}
|\varphi^C\rangle|\Phi_+^{AB}\rangle &= \frac{1}{\sqrt{2}}\big(a|0^C\rangle + b|1^C\rangle\big)\big(|0^A\rangle|0^B\rangle + |1^A\rangle|1^B\rangle\big) \\
&= \frac{1}{\sqrt{2}}\big(a|0^C\rangle|0^A\rangle|0^B\rangle + a|0^C\rangle|1^A\rangle|1^B\rangle \\
&\quad + \ b|1^C\rangle|0^A\rangle|0^B\rangle + b|1^C\rangle|1^A\rangle|1^B\rangle\big) \\
&= \frac{1}{2}\Big\{a\big(|\Phi_+^{CA}\rangle + |\Phi_-^{CA}\rangle\big)|0^B\rangle + a\big(|\Psi_+^{CA}\rangle + |\Psi_-^{CA}\rangle\big)|1^B\rangle \\
&\quad + \ b\big(|\Psi_+^{CA}\rangle - |\Psi_-^{CA}\rangle\big)|0^B\rangle + b(|\Phi_+^{CA}\rangle - |\Phi_-^{CA}\rangle)|1^B\rangle\Big\} \\
&= \frac{1}{2}\Big\{|\Phi_+^{CA}\rangle\big(a|0^B\rangle + b|1^B\rangle\big) + |\Psi_+^{CA}\rangle\big(a|1^B\rangle + b|0^B\rangle\big) \\
&\quad + \ |\Psi_-^{CA}\rangle\big(a|1^B\rangle - b|0^B\rangle\big) + |\Phi_-^{CA}\rangle\big(a|0^B\rangle - b|1^B\rangle\big)\Big\} \\
&= \frac{1}{2}\Big\{|\Phi_+^{CA}\rangle|\varphi^B\rangle + |\Psi_+^{CA}\rangle\sigma_1^B|\varphi^B\rangle + \\
&\quad + |\Psi_-^{CA}\rangle(-i\sigma_2^B)|\varphi^B\rangle + |\Phi_-^{CA}\rangle\sigma_3^B|\varphi^B\rangle\Big\}\,.
\end{aligned} \tag{11.3}$$

So far, we have only expanded mathematically in terms of a Bell basis in the space $\mathcal{H}_2^C \otimes \mathcal{H}_2^A$ which is accessible to Alice. This led via transformations with the Pauli operators

to the state $|\varphi^B\rangle$ in $\mathcal{H}_2^B$. Bob must still compensate this transformation by his interventions. To this end, Alice first takes action: she carries out a Bell measurement on the subsystems S^C and S^A, for example like that which we described in Sect. 9.2.4. There are four possible results of this measurement, which correspond to the states $|\Phi_+^{CA}\rangle$ to $|\Phi_-^{CA}\rangle$. Alice informs Bob of the result of her measurement. The associated two bits of classical information –more bits are not required – are transmitted by Alice e. g. by telephone to Bob. He applies the corresponding unitary transformation $\mathbb{1}^B, \sigma_1^B, i\sigma_2^B$ or σ_3^B to his subsystem. Owing to $\sigma_k^B \sigma_k^B = \mathbb{1}^B$, in this manner the state $|\varphi^B\rangle$ is always obtained. *We have thus described a procedure which prepares the system S^B in the state $|\varphi^B\rangle$ and the teleportation has thus succeeded.* The procedure is loss-free, since no quantum systems had to be eliminated.

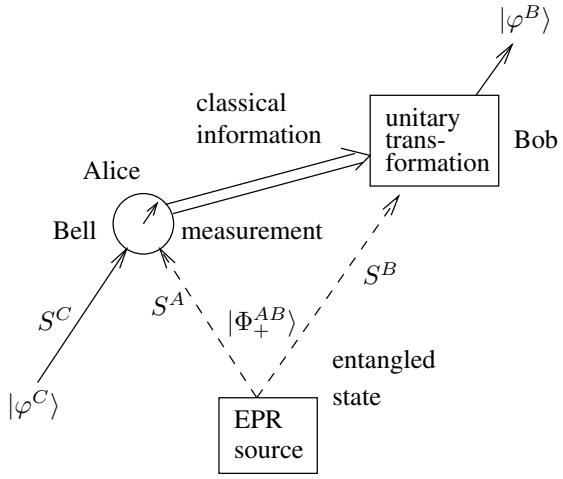


Figure 11.1: Quantum teleportation.

After the transmission of the state, none of the systems S^C and S^A are in the state $|\varphi\rangle$. This reflects the prohibition of copying. The initial entanglement of S^A and S^B was transferred to S^C and S^A. Each of the results of Alice's Bell measurement occurs with the same probability, $\frac{1}{4}$. Neither Alice nor Bob can obtain any information about the teleported state $|\varphi\rangle$ from this measurement. If the state $|\varphi\rangle$ is initially unknown, it turns up again as an unknown state $|\varphi\rangle$ in the system S^B. The special theory of relativity is not violated, since only classical information transmission was employed. The state of the subsystems S^C and S^A at Alice's location following the transmission is the completely mixed state $\frac{1}{4}\mathbb{1}^{AB}$. Although only 2 bits were transmitted, the state $|\varphi\rangle$ was exactly teleported. This however presumes that the maximally-entangled state $|\Phi_+^{AB}\rangle$ was initially present. In a practical implementation, this condition is only approximately met. We discuss in Sect. 11.6 how one can increase the degree of entanglement by means of entanglement distillation.

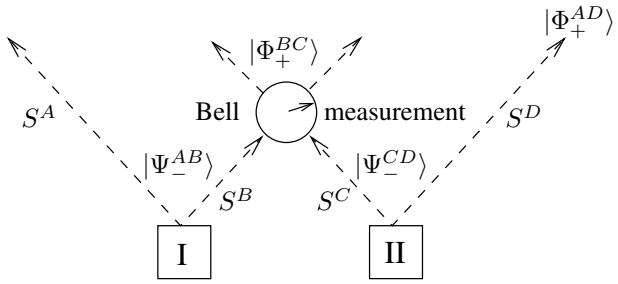


Figure 11.2: Entanglement swapping from the systems S^{AB} and S^{CD} to the systems S^{AD} and S^{BC}.

11.4 Entanglement Swapping

We described the production of entangled pairs of qubits in Sect. 8.5. In fact, it is not necessary for the entanglement of two qubits that the state be generated in a single composite process by means of unitary dynamics. Using *entanglement swapping*, two quantum systems can be placed in an entangled composite state at separate locations without any mutual interactions whatsoever. We discuss an example.

Two EPR sources I and II each produce at the same time a bipartite system S^{AB} and S^{CD} in the Bell states $|\Psi_-^{AB}\rangle$ and $|\Psi_-^{CD}\rangle$, respectively (see Fig. 11.2). All together, we thus have the product state

$$|\Psi_-^{AB}\rangle|\Psi_-^{CD}\rangle = \frac{1}{2}\left(|0^A,1^B\rangle - |1^A,0^B\rangle\right)\left(|0^C,1^D\rangle - |1^C,0^D\rangle\right) \tag{11.4}$$

in $\mathcal{H}_2^A \otimes \mathcal{H}_2^B \otimes \mathcal{H}_2^C \otimes \mathcal{H}_2^D$. We introduce Bell bases in the spaces $\mathcal{H}_2^A \otimes \mathcal{H}_2^D$ and $\mathcal{H}_2^B \otimes \mathcal{H}_2^C$. Then we can rewrite the state as the result of an intermediate computation:

$$|\Psi_-^{AB}\rangle|\Psi_-^{CD}\rangle = \frac{1}{2}\left(|\Psi_+^{AD}\rangle|\Psi_+^{BC}\rangle - |\Psi_-^{AD}\rangle|\Psi_-^{BC}\rangle - |\Phi_+^{AD}\rangle|\Phi_+^{BC}\rangle + |\Phi_-^{AD}\rangle|\Phi_-^{BC}\rangle\right) . \tag{11.5}$$

One can immediately see that a Bell measurement on the subsystems S^B and S^C transforms the previously non-entangled subsystems S^A and S^D into a Bell state. Projection, e. g. onto $|\Phi_+^{BC}\rangle$, yields $|\Phi_+^{AD}\rangle$. The Bell states of the now entangled subsystems S^{AD} and S^{BC} are, following the measurement, each the same. In this process, we are not dealing with the teleportation of states, but rather with the *transfer of entanglement*. The Bell measurement plays a similar role here as in teleportation.

It is helpful to make it once more operationally clear just how the subsystem S^{AD} is transformed into the Bell state. EPR source I is supposed to be at Alice's location and source II at Bob's location. The Bell measurement is carried out by Eve. The sources I and II repeatedly and synchronously prepare pairs of spin-$\frac{1}{2}$ particles. Alice and Bob each keep one particle and send the second one to Eve. Eve carries out a Bell measurement on each pair which is sent to her. If the output of the measurement result shows that the state $|\Phi_+^{BC}\rangle$ is present, Eve informs Alice and Bob. Only in this case do Alice and Bob send on their particles as shown

in Fig. 11.2. All the other particles are sorted out or annihilated. The remaining pairs are then in the entangled state $|\Phi_+^{AD}\rangle$. All these actions together represent a preparation procedure for the state $|\Phi_+^{AD}\rangle$. The essential elements of the preparation procedure are the non-local Bell measurement by Eve and the selection by Alice and Bob on the basis of the classical information which they obtain from Eve.

11.5 Spooky Action into the Past?*

We consider the following modification of the above scenario, which apparently leads to a further quantum-mechanical paradox: Alice and Bob each generate in I and in II simultaneously two entangled spin-$\frac{1}{2}$ particles in the states $|\Psi_-^{AB}\rangle$ or $|\Psi_-^{CD}\rangle$. They repeat this procedure many times and enumerate their pairs. The particle pairs S^B and S^C are sent to Alice. Alice gives these pairs the same number. The four particles are in the composite state $|\Psi_-^{AB}\rangle \otimes |\Psi_-^{CD}\rangle$. The two pairs of particles S^{AB} and S^{CD} are completely independent of each other. In contrast to the sequence described in the preceding section, in this case Alice first makes a measurement on her particles S^A of the spin polarisation $\boldsymbol{\sigma a}$ in the $\boldsymbol{a}$ direction and Bob measures the spin polarisation $\boldsymbol{\sigma d}$ in the $\boldsymbol{d}$ direction on his particles S^D. Later, as in the EPR experiments described in Chap. 10, the mean value of the products of the measured values will be computed. Therefore, Alice and Bob take note of the numbers of the particles and their associated measured values. The particles S^A and S^D are not used further. In a second step, Eve then carries out a Bell measurement on her pairs of particles and notes the numbers of the pairs for which the measurement yields the state $|\Phi_+^{BC}\rangle$. After many repetitions of this process, she informs Alice and Bob of these numbers, and they sort out the measured values which were associated with these numbers in their measurements. They then compute the expectation value of the products of the measurement results. In Sect. 10.1, we called this value the correlation coefficient ϵ (compare Eq. (10.15)).

Formally, we obtain ϵ^{AD} as the expectation value of a product of observables by extending Eq. (10.15) by the projection operator $|\Phi_+^{BC}\rangle\langle\Phi_+^{BC}|$ which describes selective measurements:

$$\epsilon^{AD} = \langle \Psi_-^{AB}, \Psi_-^{CD} | \boldsymbol{\sigma}^A \mathbf{a} \otimes \boldsymbol{\sigma}^D \mathbf{d} \otimes |\Phi_+^{BC}\rangle\langle\Phi_+^{BC}| | \Psi_-^{AB}, \Psi_-^{CD}\rangle \; . \tag{11.6}$$

With Eq. (11.5), we obtain in analogy to Eq. (10.15)

$$\epsilon^{AD} = \langle \Phi_+^{AD} | \boldsymbol{\sigma}^A \mathbf{a} \otimes \boldsymbol{\sigma}^D \mathbf{d} | \Phi_+^{AD}\rangle \; . \tag{11.7}$$

As we saw in Sect. 10.4, the measured results violate the CHSH inequality if one chooses for $\mathbf{a}$ and $\mathbf{d}$ the directions of the x and y axes and the axes rotated by $45°$ relative to them independently of each other and with the same frequencies (compare Eqs. (10.30) and (10.31)).

We summarise the results: measurements on pairs of independent particles S^A and S^D (produced separately in I and II), which were carried out before the measurements of Eve, led to results which violate the Bell inequality. We attributed violations of the Bell inequality in Chap. 10 to the presence of quantum correlations and hence to entanglement. Was the state of

*The sections marked with an asterisk * can be skipped over in a first reading.

the systems S^A and S^D entangled before the measurements by Eve? For the reduced density operator, we find with Eq. (11.5), taking the trace over the Bell basis of S^{BC}

$$\begin{aligned} \rho^{AD} &= \mathrm{tr}[|\Psi_-^{AB}, \Psi_-^{CD}\rangle\langle\Psi_-^{AB}, \Psi_-^{CD}|] \\ &= \frac{1}{4}\mathbb{1}^{AD} \,. \end{aligned} \tag{11.8}$$

As could be expected, this state is not entangled. Did Eve somehow produce entanglement retroactively? It would seem that in addition to a "spooky action at a distance" (cf. Sect. 7.5.2), there is also – caused by Eve – a "*spooky action into the past*" or a "*delayed entanglement*".

This is of course not the case. Three different measurements, whose sequencing is not important, were carried out. The measurements on S^A and S^D lead to a set of measured values. From this set, mean values were computed, and these are considered in the Bell inequality. It is found (cf. Eq. (11.8)) that the Bell inequality is not violated. Only if one selects the measurements based on the results of the Bell measurements carried out by Eve, and hence chooses a particular subensemble of pairs of measured values, do then the mean values computed from this subensemble lead to a violation of the Bell inequality. Selection is however not an action into the past. Since the Bell measurements on S^B and S^C play an important role in the scenario described above, there can be no locally-realistic model for the composite procedure. It is thus plausible that the possibility exists of violating the Bell inequality by means of a clever choice of the measured data.

11.6 Entanglement Distillation

Maximally-entangled pure states are the central tool in the processes of quantum information theory. We have already seen this in the cases of data compression, dense coding, and quantum teleportation. In a realistic experimental implementation, maximally-entangled states are not in fact prepared. Instead, mixtures are obtained. In practice, an additional point must be considered. Even if the preparation procedure were ideal, additional noiseless channels would be required in order to conserve the maximal entanglement during transmission of the subsystems to Alice and Bob. Technically available channels are however noisy, and give rise to mixed entangled states. These states are not maximally entangled and require manipulation. Procedures for obtaining purer states are called *purification*. The increase in the degree of entanglement is referred to as *entanglement distillation*. In general, both effects are obtained at the same time.

Typically, distillation is carried out as follows: Alice and Bob receive for example the corresponding subsystems of a non-optimally entangled composite system via noisy channels. This is repeated with a number of pairs. Alice and Bob then apply a *distillation protocol* to the subsystems, which consists of local operations and classical communication (LOCC). The local operations can be *unilateral* and act on only one subsystem, or in particular they can be *bilateral* and affect two subsystems which belong to different entangled pairs (compare the example in Fig. 11.3). In this process, usually some of the entangled pairs are sacrificed in order to enhance the entanglement of the remaining pairs. The selection accompanying this process requires a classical communication between Alice and Bob. The procedure is finally repeated with the remaining pairs, which already have a higher degree of entanglement. In the

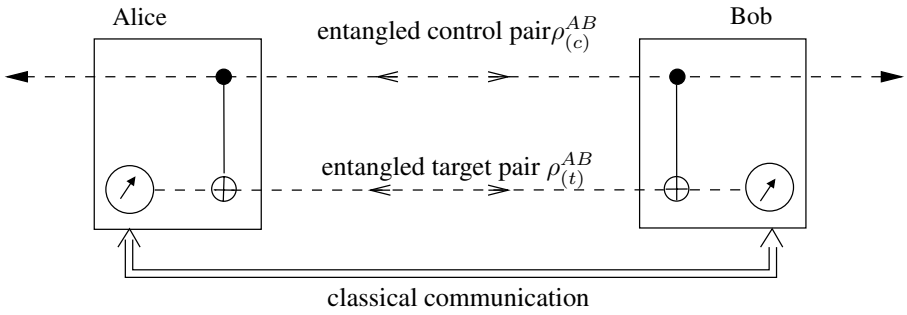


Figure 11.3: Quantum circuit for entanglement distillation. The subsystems of the control pair move off to the left and to the right (arrows). The same holds for the target pair. Following the selection, the output control pair is more strongly entangled.

following, we will discuss an example in which the typical steps of the procedure will become clear. Owing to the particular initial state chosen, this example is not very relevant in practice, but it has the advantage that all the computational steps can be readily followed.

Distillation using CNOT operations The effect of the noisy channel is supposed to consist of a mixture of the initially present, maximally-entangled Bell state $|\Phi_+^{AB}\rangle$ with the Bell state $|\Psi_+^{AB}\rangle$. *We leave off the index* $+$ *and write* $|\Phi^{AB}\rangle$ *or* $|\Psi^{AB}\rangle$. The distillation is based on two pairs of qubits which we denote by indices (c) and (t) and refer to as the *control pair* and the *target pair*. Their states are the mixtures

$$\rho_{(c)}^{AB} = F|\Phi_{(c)}^{AB}\rangle\langle\Phi_{(c)}^{AB}| + (1-F)|\Psi_{(c)}^{AB}\rangle\langle\Psi_{(c)}^{AB}| \tag{11.9}$$

$$\rho_{(t)}^{AB} = F|\Phi_{(t)}^{AB}\rangle\langle\Phi_{(t)}^{AB}| + (1-F)|\Psi_{(t)}^{AB}\rangle\langle\Psi_{(t)}^{AB}| \,. \tag{11.10}$$

F is the *fidelity*

$$F = \langle\Phi^{AB}|\rho^{AB}|\Phi^{AB}\rangle \,. \tag{11.11}$$

It indicates how similar the state ρ^{AB} is to the maximally-entangled state $|\Phi^{AB}\rangle\langle\Phi^{AB}|$ with $F = 1$; the goal of the procedure is to approach this latter state as closely as possible. We choose F as a measure of the degree of entanglement and justify this more precisely in Sect. 11.7.

For the distillation, Alice and Bob carry out a CNOT-Transformation on the qubit pair at their disposal (compare Fig. 11.3). The effect of this transformation can be described as follows: we can imagine that the mixture ρ^{AB} was prepared in such a way that the state $|\Phi^{AB}\rangle$ occurs with the probability F and the state $|\Psi^{AB}\rangle$ with the probability $(1-F)$. The mixed state $\rho_{(c)}^{AB} \otimes \rho_{(t)}^{AB}$ of all four qubits is a density operator on $\mathcal{H}_4$ and contains in its ensemble decomposition the state $|\Phi_{(c)}^{AB}\rangle|\Phi_{(t)}^{AB}\rangle$ with the probability F^2, and the state $|\Psi_{(c)}^{AB}\rangle|\Psi_{(t)}^{AB}\rangle$ with the probability $(1-F)^2$, etc.

We list the states of the ensemble decomposition of $\rho_{(c)}^{AB} \otimes \rho_{(t)}^{AB}$ together with their probabilities to the left of the arrows:

$$F^2 \quad : \quad |\Phi_{(c)}^{AB}\rangle|\Phi_{(t)}^{AB}\rangle \quad \overset{CNOT}{\underset{CNOT}{\longrightarrow}} \quad |\Phi_{(c)}^{AB}\rangle|\Phi_{(t)}^{AB}\rangle \tag{11.12}$$

$$(1-F)^2 \quad : \quad |\Psi_{(c)}^{AB}\rangle|\Psi_{(t)}^{AB}\rangle \quad \overset{CNOT}{\underset{CNOT}{\longrightarrow}} \quad |\Psi_{(c)}^{AB}\rangle|\Phi_{(t)}^{AB}\rangle \tag{11.13}$$

$$F(1-F) \quad : \quad |\Phi_{(c)}^{AB}\rangle|\Psi_{(t)}^{AB}\rangle \quad \overset{CNOT}{\underset{CNOT}{\longrightarrow}} \quad |\Phi_{(c)}^{AB}\rangle|\Psi_{(t)}^{AB}\rangle \tag{11.14}$$

$$F(1-F) \quad : \quad |\Psi_{(c)}^{AB}\rangle|\Phi_{(t)}^{AB}\rangle \quad \overset{CNOT}{\underset{CNOT}{\longrightarrow}} \quad |\Psi_{(c)}^{AB}\rangle|\Psi_{(t)}^{AB}\rangle \; . \tag{11.15}$$

Alice and Bob each use a CNOT gate. The action $\overset{CNOT}{\underset{CNOT}{\longrightarrow}}$ of these two bilateral CNOT gates is shown on the right-hand side.

To clarify the structure of the associated computations, we prove the rearrangement (11.15). To do so, we apply the rules for the CNOT gate from Chap. 7 to the states at Alice's and Bob's locations:

$$\begin{aligned} 2|\Psi_{(c)}^{AB}\rangle|\Phi_{(t)}^{AB}\rangle \quad &= \quad \Big(|0_{(c)}^{A},1_{(c)}^{B}\rangle + |1_{(c)}^{A},0_{(c)}^{B}\rangle\Big)\Big(|0_{(t)}^{A},0_{(t)}^{B}\rangle + |1_{(t)}^{A},1_{(t)}^{B}\rangle\Big) \\ &\overset{CNOT}{\underset{CNOT}{\longrightarrow}} \quad |0_{(c)}^{A},1_{(c)}^{B}\rangle\Big(|0_{(t)}^{A},1_{(t)}^{B}\rangle + |1_{(t)}^{A},0_{(t)}^{B}\rangle\Big) \\ &\quad\; + |1_{(c)}^{A},0_{(c)}^{B}\rangle\Big(|1_{(t)}^{A},0_{(t)}^{B}\rangle + |0_{(t)}^{A},1_{(t)}^{B}\rangle\Big) \\ &\quad\; = 2|\Psi_{(c)}^{AB}\rangle|\Psi_{(t)}^{AB}\rangle \; . \end{aligned} \tag{11.16}$$

We return to the equations (11.12)-(11.15). In the last step, Alice and Bob each carry out a measurement on the target pairs in the computational basis $\{|0\rangle, |1\rangle\}$ (see Fig. 11.3). Using classical communication, they inform each other of the results. If their measured values agree, the components (11.12) or (11.13) of the ensemble decomposition are at hand. The control qubit is then transferred to the state $|\Phi_{(c)}^{AB}\rangle$ with a probability F^2, or to the state $|\Psi_{(c)}^{AB}\rangle$ with the probability $(1-F)^2$. The corresponding control qubits are recycled by Alice and Bob. If the results of the measurement do not agree, they eliminate the control qubit. The resulting density operator after normalisation takes the form

$$\rho'^{AB}_{(c)} = F'|\Phi_{(c)}^{AB}\rangle\langle\Phi_{(c)}^{AB}| + (1-F')|\Psi_{(c)}^{AB}\rangle\langle\Psi_{(c)}^{AB}| \tag{11.17}$$

with

$$F' = \frac{F^2}{F^2 + (1-F)^2} \, . \tag{11.18}$$

If the initial state already had an entanglement degree of $F > \frac{1}{2}$, then it is found that $F' > F$. The degree of entanglement has been increased, as intended.

Alice and Bob can then apply the distillation protocol to two pairs in the state state $\rho'^{AB}_{(c)}$, in order to obtain a still higher degree of entanglement. If they repeat the distillation many times in this way, they can approach the maximally-entangled state $|\Psi_{+}^{AB}\rangle$ arbitrarily closely. The price for this increasing entanglement is the need to use more and more pairs and to discard many pairs. In Sect. 13.3.6, we will describe another procedure for entanglement distillation.

11.7 A Measure of Entanglement for Mixtures: Entanglement of Formation and Concurrence*

In the previous section, we took the parameter F to be a measure of the entanglement of the mixed state $\rho_{(c)}^{AB}$ of Eq. (11.9). We wish to justify that choice in the following. For this, we give a general measure of the entanglement of qubit systems, which is also applicable to density operators. We start from the degree of entanglement for pure states. To keep this section short, we dispense with presenting longer but simple computations in detail and refer to the literature for the central lemma.

Pure states We consider systems S^{AB} with two subsystems S^A and S^B. Within this entire section, we assume that the subsystems are qubits. For pure states $|\psi^{AB}\rangle$, we encountered the entropy of the subsystems S^A or S^B as a measure $E(\psi^{AB})$ of entanglement in Sect. 8.3.3:

$$E(\psi^{AB}) = S(A) = -\text{tr}[\rho_A \log \rho_A] \tag{11.19}$$

$$= S(B) = -\text{tr}[\rho_B \log \rho_B] \; . \tag{11.20}$$

Here, ρ^A and ρ^B are the reduced density operators of the subsystems S^A and S^B.

We expand the state $|\psi^{AB}\rangle$ in terms of the computational basis

$$|\psi^{AB}\rangle = a|0^A, 0^B\rangle + b|0^A, 1^B\rangle + c|1^A, 0^B\rangle + d|1^A, 1^B\rangle \tag{11.21}$$

and evaluate Eq. (11.19). A lengthy intermediate computation, which we shall not reproduce here, leads to

$$E(\psi^{AB}) = h\left(\frac{1+\sqrt{1-C(\psi^{AB})^2}}{2}\right) \; . \tag{11.22}$$

Here, we have introduced the binary entropy function

$$h(x) := -[x \log_2 x + (1-x)\log_2(1-x)] \tag{11.23}$$

and the abbreviation

$$C(\psi^{AB}) := 2|ad - bc| \; , \; C \leq 1 \; . \tag{11.24}$$

$C(\psi^{AB})$ is called the *concurrence* of the state $|\psi^{AB}\rangle$. The entanglement E is a monotonic function of C. Both have a range of values from 0 to 1 and are identical at the endpoints of this range. The concurrence can therefore likewise be considered to be a measure of entanglement. A state with $E = C = 0$ is *separable*. When $E = C = 1$, the state is *maximally entangled*.

As a preparation for the generalisation to mixtures, we note that the concurrence can also be written in the form

$$C(\psi^{AB}) = |\langle \psi^{AB} | \bar{\psi}^{AB} \rangle| \; . \tag{11.25}$$

*The sections marked with an asterisk * can be skipped over in a first reading.

$|\bar{\psi}^{AB}\rangle$ is obtained here from $|\psi^{AB}\rangle$, by first taking the complex conjugate prefactors in the computational basis, that is in Eq. (11.21), ($a \to a^*$, $b \to b^*$, etc.). The result will be denoted by $|\psi^{AB}\rangle^*$. Then the operator $\sigma_y^A \otimes \sigma_y^B$ is applied to this state:

$$|\bar{\psi}^{AB}\rangle := \sigma_y^A \otimes \sigma_y^B |\psi^{AB}\rangle^* \,. \tag{11.26}$$

σ_y exchanges the states of the computational basis and inserts the relative phase $\pm i$. Equation (11.26) leads with Eq. (11.25) to Eq. (11.24).

We also write the concurrence of a pure state as a function of the associated density operator $\rho^{AB} = |\psi^{AB}\rangle\langle\psi^{AB}|$:

$$\begin{aligned} C(\psi^{AB})^2 &= |\langle\psi^{AB}|\bar{\psi}^{AB}\rangle|^2 = \langle\psi^{AB}|\bar{\psi}^{AB}\rangle\langle\bar{\psi}^{AB}|\psi^{AB}\rangle \\ &= \mathrm{tr}[|\psi^{AB}\rangle\langle\psi^{AB}|\bar{\psi^{AB}}\rangle\langle\bar{\psi}^{AB}|] \\ &= \mathrm{tr}[\rho^{AB}\bar{\rho}^{AB}] = \\ &= \mathrm{tr}[R^{AB}] \,. \end{aligned} \tag{11.27}$$

The operator

$$R^{AB} := \rho^{AB}\bar{\rho}^{AB} \tag{11.28}$$

is obtained from

$$\bar{\rho}^{AB} = |\bar{\psi}^{AB}\rangle\langle\bar{\psi}^{AB}| = \left(\sigma_y^A \otimes \sigma_y^B\right) \rho^{*AB} \left(\sigma_y^A \otimes \sigma_y^B\right) \,. \tag{11.29}$$

ρ^{*AB} is generated by writing ρ^{AB} as a matrix in terms of the computational basis and transforming to the complex conjugate matrix elements.

Mixtures With this representation, an initially purely formal generalisation of the concurrence to mixtures ρ^{AB}, which do not describe pure states, suggests itself. We adopt Eqs. (11.27) through (11.29) and compute

$$E(\rho^{AB}) = h\left(\frac{1+\sqrt{1-C(\rho^{AB})^2}}{2}\right) \,. \tag{11.30}$$

To evaluate $C(\rho^{AB})$ for density operators, we employ a *lemma of Wooters [Woo 98], according to which the concurrence can be written explicitly in the form*

$$C(\rho^{AB}) = \max\{0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4\} \,. \tag{11.31}$$

The λ_1, λ_2, λ_3 and λ_4 are here the square roots of the eigenvalues of the matrix $R^{AB} = \rho^{AB}\bar{\rho}^{AB}$. They are non-negative real numbers and λ_1 is the largest eigenvalue. The λ_i are at the same time the eigenvalues of the matrix $\sqrt{\sqrt{\rho^{AB}}\bar{\rho}^{AB}\sqrt{\rho^{AB}}}$.

Entanglement of formation The question arises as to whether the quantity $E(\rho^{AB})$ is only a mathematical generalisation or also represents a physically reasonable measure of the entanglement for mixtures. Every density operator ρ^{AB} can be represented as the density operator of an ensemble of pure states $|\psi_l\rangle$:

$$\rho^{AB} = \sum_l p_l |\psi_l^{AB}\rangle\langle\psi_l^{AB}| \,. \tag{11.32}$$

The mean entanglement of the state ρ^{AB} is found as the mean value of the entanglement $E(\psi_l^{AB})$ of the pure states in the form $\sum_l p_l E(\psi_l)$. The *entanglement of formation* of ρ^{AB} is defined as the minimum value of the mean entanglement if one considers all the possible ensemble decompositions of ρ^{AB}:

$$E_f(\rho^{AB}) = \min \sum_l p_l E(\psi_l^{AB}) \,. \tag{11.33}$$

If one thus prepares the state ρ^{AB} as a statistical mixture, then at least the entanglement $E_f(\rho^{AB})$ must be produced on the average.

The connection to the preceding considerations is finally established by means of the theorem of Wooters [Woo 98]: *For a system of two qubits which is in the state ρ^{AB}, the entanglement of formation $E_f(\rho^{AB})$ as a function of the concurrence $C(\rho^{AB})$ is given by*

$$E_f(\rho^{AB}) = E(\rho^{AB}) \tag{11.34}$$

with $E(\rho^{AB})$ from Eq. (11.30). Therefore, $E(\rho^{AB})$ can be interpreted operationally as a physically reasonable measure of the entanglement for mixed states as well.

The magic basis We give yet another alternative procedure for computing the concurrence $C(\rho^{AB})$. Instead of making use of the computational basis as in the above considerations, it may be expedient to begin with the *magic basis*. It consists of Bell states with particular phases:

$$\begin{array}{ll} |e_1^{AB}\rangle := |\Phi_+^{AB}\rangle \,, & |e_2^{AB}\rangle := i|\Phi_-^{AB}\rangle \,, \\ |e_3^{AB}\rangle := i|\Psi_+^{AB}\rangle \,, & |e_4^{AB}\rangle := |\Psi_-^{AB}\rangle, \,. \end{array} \tag{11.35}$$

We first expand the pure state $|\psi^{AB}\rangle$ in terms of this basis

$$|\psi^{AB}\rangle = \sum_i \alpha_i |e_i^{AB}\rangle \tag{11.36}$$

($\sum_i |\alpha_i|^2 = 1$) and take the complex conjugate state

$$|\psi^{AB}\rangle_* := \sum_i \alpha_i^* |e_i^{AB}\rangle \,. \tag{11.37}$$

One can readily confirm that this result agrees with the vector $|\bar{\psi}^{AB}\rangle$ from Eq. (11.26)

$$|\psi^{AB}\rangle_* = |\bar{\psi}^{AB}\rangle \,. \tag{11.38}$$

For the concurrence of the pure state $|\psi^{AB}\rangle$, we obtain with Eq. (11.25) the simple expression

$$C(\psi^{AB}) = \left| \sum_i \alpha_i^2 \right| . \tag{11.39}$$

Note that the complex numbers α_i, and not their magnitudes, are squared. A pure state is for example maximally entangled ($C = 1$, $E = 1$) just when the α_i all have the same phase ($\alpha_i = e^{i\varphi}|\alpha_i|$).

With equation (11.38), the results of Eq. (11.27) and (11.28) can be applied to density operators

$$C^2(\rho^{AB}) = \mathrm{tr}[R^{AB}] \tag{11.40}$$

$$R^{AB} = \rho^{AB}\rho_*^{AB} . \tag{11.41}$$

Here, ρ_*^{AB} for the pure state $|\psi^{AB}\rangle$ is

$$\rho_*^{AB} = |\psi^{AB}\rangle_{**}\langle\psi^{AB}| = \sum_{i,j} |e_i^{AB}\rangle\langle e_j^{AB}|\psi^{AB}\rangle\langle\psi^{AB}|e_i^{AB}\rangle\langle e_j^{AB}| \tag{11.42}$$

and analogously, for the mixture ρ^{AB}, we have

$$\rho_*^{AB} = \sum_{i,j} |e_i^{AB}\rangle\langle e_j^{AB}|\rho^{AB}|e_i^{AB}\rangle\langle e_j^{AB}| . \tag{11.43}$$

We made use of Eq. (11.37) for the rearrangement. In order to obtain ρ_*^{AB}, one can write ρ^{AB} in terms of the magic basis as a matrix and take the complex conjugates of the matrix elements.

The direct computation of the concurrence $C(\rho^{AB})$ can be carried out with the aid of Eq. (11.31) by again making use of the lemmas. The λ_i are the eigenvalues of $R^{AB}(\rho)$. We can however now evaluate R^{AB} in the form of Eq. (11.41). This is especially favourable when ρ^{AB} is given in the Bell basis. If it is given in the computational basis, it is conversely more simple to make use of the analogous equation (11.28).

An example In terms of the magic basis, the concurrence $C(F)$ for the state ρ^{AB} of Eq. (11.9) from the previous section is particularly simple to compute. With

$$\rho^{AB} = \rho_*^{AB} = \sqrt{R^{AB}} = \begin{pmatrix} F & & & 0 \\ & 0 & & \\ & & 1-F & \\ 0 & & & 0 \end{pmatrix} , \tag{11.44}$$

we obtain by evaluation with the aid of Eq. (11.31) for the dependence of the concurrence on F

$$0 \le F \le \frac{1}{2} \quad : \quad C(F) = 1 - 2F \tag{11.45}$$

$$\frac{1}{2} \le F \le 1 \quad : \quad C(F) = 2F - 1 . \tag{11.46}$$

For $F \ge \frac{1}{2}$, the concurrence increases linearly with F. We thus justifiably employed F in the previous section as a measure of the entanglement.

11.8 Complementary Topics and Further Reading

- An introductory overview of entanglement as a tool: [Wer 06].
- Quantum cryptography; the classics: [Eke 91], [BBM 92].
- In-depth literature on defending against eavesdropping: [HN 99], [Gru 99], [GRT 02], [Lom 02a].
- Protocols with more than two states: [BB 84], [GRT 02], [Lom 02a].
- The B92 protocol: [HAD 95], [Gru 99], [Lom 02a].
- Experimental quantum cryptography: [HAD 95], [Zbi 98], [HN 99], [BD 00], [EGH 00], [GM 02], [GRT 02].
- BBM92: [BBM 92], [GM 02, p. 369].
- Dense coding; the classic: [BW 92]. Experiments: [BHL 02], [BEZ 00, p. 62], [BD 00].
- Teleportation; the classic: [BBC 93].
- Teleportation of higher-dimensional states and mixtures [Wer 01, p. 53], [Key 02, p. 474], [vLo 02, 1232].
- Experiments on teleportation: [BD 00], [BEZ 00], [BHL 02], [GM 02, p. 363].
- Entanglement swapping: [ZZH 93], [BVK 98].
- Entanglement distillation: [BBP 96], [BVS 96]
- In [BeB 96], it is shown how by applying entanglement distillation to an arbitrary mixture ρ^{AB}, a mixture ρ'^{AB} with $F' = \langle \Psi_-^{AB} | \rho'^{AB} | \Psi_-^{AB} \rangle > F$ can be prepared with the fidelity $F = \langle \Psi_-^{AB} | \rho^{AB} | \Psi_-^{AB} \rangle > \frac{1}{2}$. By means of iteration, the resulting state can be made to approach the Bell state $|\Psi_-^{AB}\rangle$ arbitrarily closely.
- On entanglement of formation and concurrence: see [HW 97] and [Woo 98] as well as the review article [Woo 01].
- Concurrence as a measure of entanglement for systems with more than two parts and states in higher-dimensional Hilbert spaces: [MCK 05].

11.9 Problems for Chapter 11

Prob. 11.1 Counterfeit-proof banknotes.
Assume that it is technically possible to store photons on a banknote in individual cells over a long period of time. How can one use this to print banknotes which are secure against counterfeiting?

Prob. 11.2 [for 11.3]: What modifications of the considerations in Sect. 11.3 would be required if the state $|\Psi_-^{AB}\rangle = \frac{1}{\sqrt{2}}\left(|0,1\rangle - |1,0\rangle\right)$ were used instead of $|\Phi_+^{AB}\rangle$?

Prob. 11.3 [for 11.1]: Assume that the eavesdropper carries out a polarisation measurement on each of the two photons (spin-$\frac{1}{2}$ particles) and then lets the photons pass on to Alice or Bob. The eavesdropper thus prepares pairs of photons with a probability $p(\theta^A, \theta^B)$ with photons in the states $|\theta^A\rangle$ and $|\theta^B\rangle$. The two polarisation directions are described by the angles θ^A and θ^B. Determine the correlation coefficients explicitly and show that $|S| \leq \sqrt{2}$ holds. Alice and Bob would therefore be able to detect this eavesdropping.

Prob. 11.4 [for 11.3]: In which state is the system S^B before Alice carries out her measurements? If Alice performs Bell measurements on S^C and S^A but does not inform Bob of this, in which state is the system S^B?

Prob. 11.5 [for 11.7]: Prove Eq. (11.25) and the assertion $C \leq 1$ from Eq. (11.24).

Prob. 11.6 [for 11.7]: Give a visualisation of ρ^A and $\bar{\rho}^A = \sigma_y^A \rho^{*A} \sigma_y^A$ making use of the Bloch sphere.

Prob. 11.7 [for 11.7]: Give the equation for ρ^{*AB} which is analogous to Eq. (11.43).

Prob. 11.8 [for 11.7]: A mixture of Bell states

$$\rho^{AB} = \lambda_1 |\Phi_+^{AB}\rangle\langle\Phi_+^{AB}| + \lambda_2 |\Phi_-^{AB}\rangle\langle\Phi_-^{AB}| + \lambda_3 |\Psi_+^{AB}\rangle\langle\Psi_+^{AB}| + \lambda_4 |\Psi_-^{AB}\rangle\langle\Psi_-^{AB}| \quad (11.47)$$

with $0 \leq \lambda_1 \leq 1$, $\sum_i \lambda_i = 1$ is called a *Bell diagonal state*. Show that ρ is entangled if and only if $\max \lambda_j > \frac{1}{2}$.

© 2007 WILEY-VCH Verlag GmbH & Co.

12 The Quantum Computer

In this chapter, we wish to demonstrate how computations can be carried out using quantum systems and how such computations differ from those performed in a classical computer. It should become clear which types of problems can be treated more effectively using a quantum computer. This superiority can be traced back to the typical non-classical structures of quantum mechanics such as superposition and entanglement.

12.1 Registers and Networks

Registers A sequence of n qubit systems represents a *quantum register* of length n. Each qubit acts as a digit in the register. The state of the register is described by a vector $|\psi^{\text{in}}\rangle$ in the Hilbert space $\mathcal{H}_2^{\otimes n} = \mathcal{H}_2 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_2$ with n factor spaces $\mathcal{H}_2$. We will work in all the Hilbert spaces using the computational basis $\{|0\rangle, |1\rangle\}$. In the registers, the information is stored in binary form. The natural number a is associated with the register state

$$|a\rangle = |a_{n-1}\rangle|a_{n-2}\rangle \ldots |a_0\rangle\,, \quad a_i \in \{0,1\} \tag{12.1}$$

in $\mathcal{H}_2^{\otimes n}$. As with a classical computer, we employ the binary representation of a,

$$a = a_{n-1}2^{n-1} + a_{n-2}2^{n-2} + \cdots + a_0 2^0 \leftrightarrow (a_{n-1}, a_{n-2}, \ldots, a_0)\,. \tag{12.2}$$

There are $d := 2^n$ states of this type. They form the orthonormal computational basis of $\mathcal{H}_2^{\otimes n}$. The natural numbers $a = 0$ to $a = d-1$ enumerate the basis states. With $a \in \{0,1\}^n$, one denotes the fact that the state $|a\rangle$ is an element of the computational basis for a register of length n. For example, $6 \in \{0,1\}^3$ and the associated state has the form $|6\rangle = |1,1,0\rangle$.

It is an important property of a quantum register that it can store several numbers at the same time in orthogonal and thus by suitable measurements distinguishable states. An example is provided by

$$\frac{1}{\sqrt{2}}\left(|0,1,1\rangle + |1,1,1\rangle\right) = \frac{1}{\sqrt{2}}\left(|3\rangle + |7\rangle\right)\,. \tag{12.3}$$

The general state of a register of length n is a linear combination of the basis states

$$|\psi\rangle = \sum_{a=0}^{d-1} c_a |a\rangle\,, \quad \sum_{a=0}^{d-1} |c_a|^2 = 1\,. \tag{12.4}$$

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

The number d of basis states which can be stored simultaneously increases exponentially with the length n of the register. For $n = 270$, the number is already greater than the estimated number of atoms in the universe. This enormous storage and processing capacity is one of the strengths of the quantum computer.

This strength can however be nullified by the peculiarities of quantum measurements if they are carried out ineptly. It must be taken into account that measurements in the computational basis on the output state $|\psi^{\text{out}}\rangle$, which are performed sequentially or simultaneously on the register digits, always lead to one of the associated basis states. They allow only a single number a to be read out. In the case of a measurement on the state (12.3), that is the number 3 or 7. Such measurements in the computational basis are called *bit by bit measurements* A subsequent measurement on the resulting state yields no further information.

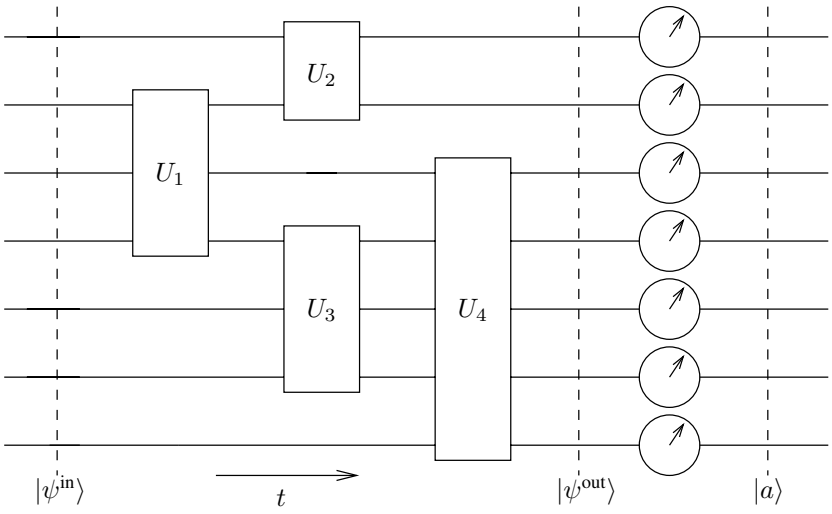


Figure 12.1: The schematic of a simple quantum network with a bit by bit measurement.

Networks The manipulations of the register states by the quantum computer are carried out using unitary transformations on $\mathcal{H}_2^{\otimes n}$. A *quantum gate* carries out a well-defined unitary transformation, which often is analogous to a logic gate in a classical computer. A *quantum network* or a *quantum circuit* consists of a number of quantum gates which act on the state in a temporally ordered fashion, either sequentially or simultaneously (compare Fig. 12.1). The gates are connected to one another via *quantum wires*, which are associated with one of the subspaces $\mathcal{H}_2$ of $\mathcal{H}_2^{\otimes n}$ and hence with one of the quantum systems. Ideal wires do not modify the state. Real wires are usually sources of errors. We have already encountered such gates and networks in Chaps. 3 and 7.

A *quantum computer* is a quantum network which carries out a *quantum computation* by transforming an input state $|\psi^{\text{in}}\rangle$ in a unitary fashion into an output state $|\psi^{\text{out}}\rangle$. Measurements are performed as a rule projectively and bit by bit upon one or all of the qubits (register digits)

of the input state. In the most general case, the unitary evolution can also be interrupted by measurements in one or several factor spaces $\mathcal{H}_2$.

For the experimental implementation of quantum networks, it is important that well-defined unitary transformations can be induced separately in a controlled manner on individual factor spaces or on products of factor spaces. These unitary transformations of the subsystem by the quantum gates take place via an external manipulation of the qubit systems or by interactions with neighbouring qubits. Their implementation is one of the major challenges to the construction of quantum computers.

12.2 Functional Computation

For the calculation of a *Boolean function* $f : \{0,1\}^n \to \{0,1\}^m$ using a quantum computer, a first register of length n is required, into which the input state $|x\rangle$ is read in, as well as a second register of length m in which the output state of the functional value $f(x)$ is stored. Both registers have finite lengths. The computation is therefore carried out according to the rules of *modular arithmetic*. These describe calculations with *remainders*. The symbol $a \bmod N$ denotes the remainder which occurs in the division of the natural number a by the natural number N[1]. Therefore, $a = qN + r$ with $q \in \mathbb{N}$ and $r = a \bmod N$ holds. Equations which contain $\bmod N$ on their right-hand sides describe the equality of the remainders (e. g. $1 = 9 = 25 \bmod 8$). For our calculations with remainders, we will in the main require addition. Here, we have

$$(a + b) \bmod N = ((a \bmod N) + (b \bmod N)) \bmod N. \tag{12.5}$$

One often leaves off the specification $\bmod N$ in the case of addition and writes simply

$$(a + b) \bmod N =: a \oplus b\ . \tag{12.6}$$

Two registers Quantum computers are based upon unitary and hence reversible state evolutions. Functions f, which do not allow a one-to-one mapping (for which therefore $f(x_1) = f(x_2)$ for arguments $x_1 \neq x_2$ holds), cannot be directly calculated by means of unitary operations (cf. Fig. 12.2). This problem is solved by carrying the argument x in a first register in unchanged form. *Thus, it is necessary to have two registers*, a first register (x register) of length n and a second register (y register) of length m. The unitary transformation for the determination of $f(x)$ then acts upon a state $|x, y\rangle$ of $\mathcal{H}_2^{\otimes n} \otimes \mathcal{H}_2^{\otimes m}$ in the following manner:

$$|x, y\rangle \xrightarrow{U_f} |x, (y + f(x)) \bmod 2^m\rangle = |x, y \oplus f(x)\rangle\ . \tag{12.7}$$

Figure. 12.3 shows this schematically. U_f is a controlled operation, since what happens to the content of the second registers depends on the content of the first. The CNOT gate which we described in Sect. 7.8 is a special case with $m = n = 1$ and $f(x) = x$. We make use of the same graphic representation as in that section. In the general case, U_f as in Eq. (12.7) is obtained from the circuit shown in Fig. 12.4.

[1] *Integers* $\mathbb{Z}$: $\{\ldots, -2, -1, 0, +1, +2, \ldots\}$. *Natural numbers* $\mathbb{N}$: $\{0, 1, 2, \ldots\}$. They are also called positive integers or non-negative integers.

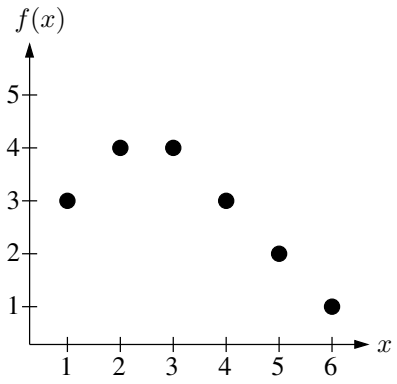


Figure 12.2: A non-uniquely reversible function.

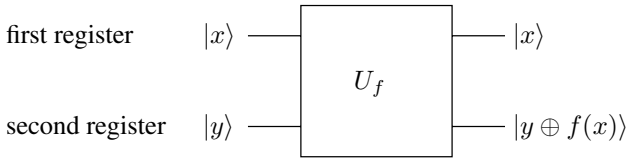


Figure 12.3: Calculation of a function as a unitary transformation. Here, its action on the vectors of the computational basis $|x\rangle \in \mathcal{H}_2^{\otimes n}$ and $|y\rangle \in \mathcal{H}_2^{\otimes m}$ is shown.

We give an example to illustrate Eq. (12.7). The quantum registers are assumed to have the lengths $n = 2$ and $m = 3$. The Boolean function f is then of the form $f : \{0,1\}^2 \to \{0,1\}^3$. We consider in particular the computation of the function $f(x) = x^2$

$$|x, 0\rangle \xrightarrow{U_f} |x, x^2 \bmod 2^3\rangle \; . \tag{12.8}$$

The unitary transformation U_f must then perform the following actions:

$$\begin{aligned} |0,0\rangle|0,0,0\rangle &\to |0,0\rangle|0,0,0\rangle \\ |0,1\rangle|0,0,0\rangle &\to |0,1\rangle|0,0,1\rangle \\ |1,0\rangle|0,0,0\rangle &\to |1,0\rangle|1,0,0\rangle \\ |1,1\rangle|0,0,0\rangle &\to |1,1\rangle|0,0,1\rangle \; . \end{aligned} \tag{12.9}$$

Here, we have used $9 \bmod 2^3 = 1$ and have written out the vectors from $\mathcal{H}_2^{\otimes 2}$ and $\mathcal{H}_2^{\otimes 3}$ as product vectors.

It remains to answer the question as to whether such unitary transformations U_f can always be implemented by using gates. It can be shown that for every Boolean function $f : \{0,1\}^n \to \{0,1\}^m$, the quantum network which carries out the transformation U_f and thus permits the computation of every function f by the quantum computer can be constructed solely of Toffoli gates. This also guarantees the unitarity of U_f. The Toffoli gate is in this sense a universal reversible gate. For the proof of this statement, we refer to the literature (see Sect. 8.8).

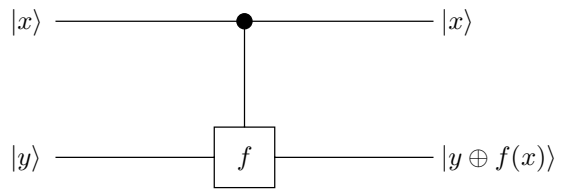


Figure 12.4: Computation of functions as a controlled operation. An alternative representation of the circuit in Fig. 12.3

Unitarity We wish to examine the simple special case $n = m = 1$ of Eq. (12.7) and to show that for every function $f : \{0,1\} \to \{0,1\}$, the transformation U_f which carries out the addition is a unitary transformation on $\mathcal{H}_2 \otimes \mathcal{H}_2$. It can thus be implemented by means of a combination of simple quantum gates. We find

$$U_f U_f |x,y\rangle = U_f |x, y \oplus f(x)\rangle = |x, y \oplus f(x) \oplus f(x)\rangle = |x,y\rangle \ , \tag{12.10}$$

and thus $U_f U_f = \mathbb{1}$.

It still remains to show that $U_f^\dagger = U_f$ holds. There are four functions f_i:

$$\begin{array}{llll} f_1(0) & = 0 \ , & f_1(1) & = 0 \\ f_2(0) & = 1 \ , & f_2(1) & = 1 \\ f_3(0) & = 0 \ , & f_3(1) & = 1 \\ f_4(0) & = 1 \ , & f_4(1) & = 0 \ . \end{array} \tag{12.11}$$

We investigate them with reference to the matrix U_f obtained with the computational basis. For f_1, we have $U_f|x,y\rangle = U_f|x, y \oplus 0\rangle = |x,y\rangle$ and hence $U_f = \mathbb{1} = U_f^\dagger$. For f_2, the equations $U_f|0,0\rangle = |0,1\rangle$, $U_f|0,1\rangle = |0,0\rangle$, $U_f|1,0\rangle = |1,1\rangle$ and $U_f|1,1\rangle = |1,0\rangle$ apply. To find the matrix representation of U_f, we can read off from them using Eq. (8.49)

$$U_f = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = U_f^\dagger \ . \tag{12.12}$$

The unitarity of U_f for f_3 and f_4 can be demonstrated in an analogous manner.

Kick back We append an observation which will prove to be useful in the following sections: we consider a function $f : \{0,1\}^n \to \{0,1\}$, i. e. the case $m = 1$, and generate the superposition $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ with $|0\rangle, |1\rangle \in \mathcal{H}_2$ as the initial state in the register. The effect of U_f then consists of

$$\begin{aligned} |x\rangle \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \quad &\xrightarrow{U_f} \quad |x\rangle \frac{1}{\sqrt{2}} (|f(x)\rangle - |1 \oplus f(x)\rangle) \\ &= \quad |x\rangle (-1)^{f(x)} \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \ . \end{aligned} \tag{12.13}$$

Here, we have made use of the fact that $f(x)$ can be only 0 or 1. Basis vectors $|x\rangle$ which give $f(x) = 1$ are multiplied by -1. The argument x thus controls a sign flip. A relative phase with respect to a superposition in the first register can be generated (see Fig. 12.5):

$$(c_1|x_1\rangle + c_2|x_2\rangle)\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \xrightarrow{U_f} [(-1)^{f(x_1)}c_1|x_1\rangle + (-1)^{f(x_2)}|x_2\rangle]\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \; . \tag{12.14}$$

Although the computation of the function and the addition take place in the second register, the state in the second register remains unchanged, $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, and sign reversals occur only in the first register, depending on $f(x)$. This process, which we will use repeatedly, is called a *kick back*.

12.3 Quantum Parallelism

By parallel application of Hadamard gates on the "binary expressed" register state $|0, 0, \ldots, 0\rangle$ of the first register of length n ($d := 2^n$),

$$\begin{aligned} H^{\otimes n}|0, 0, \ldots, 0\rangle &= \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)\frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)\ldots\frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \\ &= \frac{1}{\sqrt{d}}\sum_{x=0}^{d-1}|x\rangle =: |\Omega\rangle \, , \end{aligned} \tag{12.15}$$

a uniformly-weighted superposition of the d vectors of the computational basis of $\mathcal{H}_2^{\otimes n}$ is formed. We can consider the state $|\Omega\rangle$ to be the "superposition" of the numbers $0 \leq x \leq d-1, x \in \mathbb{N}$. If we add on the second register of length m in the state $|0\rangle \in \mathcal{H}_2^{\otimes m}$ and allow the unitary transformation to act upon $|\psi\rangle = |\Omega\rangle|0\rangle$, it is transformed into the state

$$\begin{aligned} |\psi'\rangle &= U_f\left(\frac{1}{\sqrt{d}}\sum_{x=0}^{d-1}|x, 0\rangle\right) \\ &= \frac{1}{\sqrt{d}}\sum_{x=0}^{d-1}U_f\left(|x, 0\rangle\right) \\ &= \frac{1}{\sqrt{d}}\sum_{x=0}^{d-1}|x, f(x) \bmod m\rangle \; . \end{aligned} \tag{12.16}$$

As a result of the linearity of U_f and the superposition (12.15) in the state $|\Omega\rangle$ in the first register, the value of $f(x)$ is simultaneously computed for $d = 2^n$ arguments by a single pass through the network. This parallel processing of information is termed *quantum parallelism*. d increases exponentially with the register length n. Apart from trivial cases, the resulting state $|\psi'\rangle$ is entangled.

If we measure the first n qubits in the state $|\psi'\rangle$ of Eq. (12.16) (i. e. the x register) with respect to the standard basis, we obtain one of the states $|x\rangle$ with a uniform probability $\frac{1}{d}$. If, for example, $|x_0\rangle$ results from the measurement, the composite state is transformed into

$|x_0, f(x_0)\rangle$. A measurement on the second register yields the function value $f(x_0)$. In subsequent measurements, no statements about $f(x)$ for other x values can be made. In terms of such a calculation of $f(x)$ argument by argument, the quantum computer is less efficient than a classical computer, since for the latter, the value x_0 can be set as desired. In the quantum computer, on the other hand, the measured value x_0 is a random event. The superiority of the quantum computer asserts itself for other kinds of problems.

Quantum algorithms The superiority of *quantum algorithms* as compared to classical algorithms is based in the first instance on the use of superpositions and entanglement for quite specific problems. In the main, the following two techniques are employed:

(i) *Searching for global properties* of a function $f(x)$, such as its period. For this purpose, in contrast to the classical computer, one does not first compute functional values and then compare them, but rather investigates directly the correlations between the states of the output register. We will encounter this in the case of the Deutsch problem, the Deutsch-Jozsa problem, and in the Shor algorithm.

(ii) *Amplitude amplification*, usually in an iterative manner. Here, the superposition is transformed in such a way that the state with the result being sought obtains a particularly large amplitude and will thus be measured with a high probability. As an example of this, we treat the Grover algorithm.

12.4 Two Simple Quantum Algorithms

12.4.1 The Deutsch Problem

We discuss the situation in which a function $f : \{0, 1\} \to \{0, 1\}$ is given in a *black box* or via an *oracle*. The black box can compute the function's value $f(x)$ for any input argument x. As with an oracle, one can pose a *query* and obtain an answer in each case. The problem consists of determining particular properties of $f(x)$ with a minimum number of queries. We compare a classical black box with a quantum-mechanical black box, in which $f(x)$ is implemented as a quantum algorithm. The black box carries out the transformation U_g with a well-determined function $f(x)$ which is however unknown to us.

There are four functions $f(x)$, which are listed in Eq. (12.11): The functions $f_1(x)$ and $f_2(x)$ are termed constant. The functions $f_3(x)$ and $f_4(x)$ are called balanced, since they yield an equal number of values 0 and 1. Both are global properties of the functions. It is to be determined whether the function $f(x)$ in the black box is constant or balanced. To do this, in the classical case, the computation must be carried out with the values $x = 0$ and $x = 1$. Thus, the oracle must be queried two times. In the quantum computer, we make use of the Deutsch algorithm ([Deu 85]), and do not ask, "Which functional value?", but rather directly "Which function type?". We employ the kick back of Eq. (12.14) and use the state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ in the x register. This can be done by inputting the value $|0\rangle$ there and carrying out a Hadamard transformation (compare Fig. 12.5). After applying U_f as in Eq. (12.14), the composite state

$$\frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle - |1\rangle) \xrightarrow{U_f} \frac{1}{2}\left[(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle\right](|0\rangle - |1\rangle) \qquad (12.17)$$

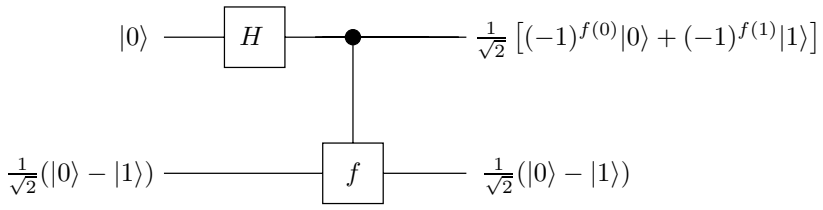


Figure 12.5: Generating a kick back in the first register.

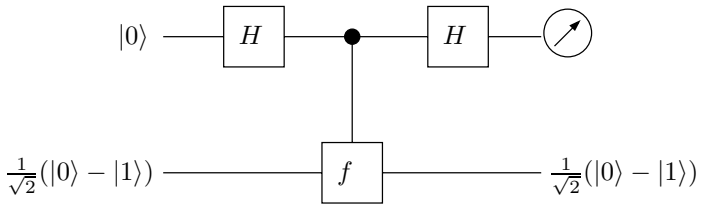


Figure 12.6: A quantum circuit for the Deutsch algorithm.

is present. The second register is not considered. If $f(x)$ is constant, the first register contains the state

$$\frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \Leftrightarrow f \text{ constant} . \tag{12.18}$$

In the balanced case, the state

$$\frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right) \Leftrightarrow f \text{ balanced} \tag{12.19}$$

is present. For the measurement, we carry out a second Hadamard transformation H as in Fig. 12.6, which leads to

$$|0\rangle \Leftrightarrow f \text{ constant} \quad \text{or} \quad |1\rangle \Leftrightarrow f \text{ balanced} . \tag{12.20}$$

A single measurement in the computational basis then yields with certainty the answer being sought.

12.4.2 The Deutsch-Jozsa Problem

We increase the length of the first register and consider functions $f : \{0,1\}^n \to \{0,1\}$, which are defined for $d = 2^n$ values of x. $f(x)$ is taken to be either constant or balanced. Balanced means that half of all the functional values are zero and the other half are one. The problem is once again to find out which type of function is contained in the black box.

The quantum algorithm of Deutsch and Jozsa ([DJ 92]) is generated by extending the Deutsch algorithm. We follow the circuit shown in Fig. 12.7. The Hadamard gates $H^{\otimes n}$ have

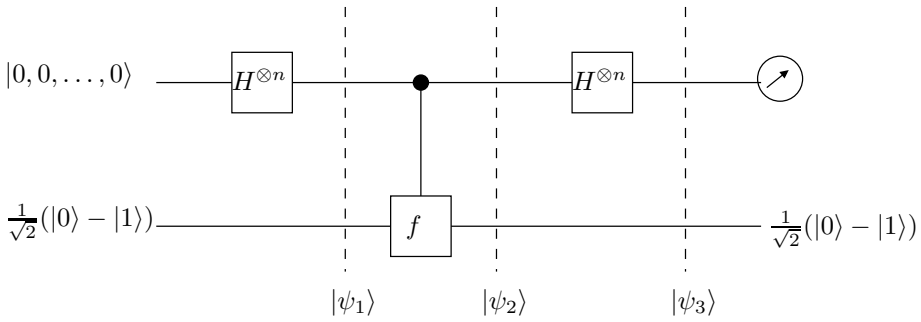


Figure 12.7: A quantum circuit for the Deutsch-Jozsa algorithm.

the effect described in Eq. (12.15) on the states of the first register.

$$|0,0,\ldots,0\rangle \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right) \xrightarrow{H^{\otimes n}\otimes \mathbb{1}} |\psi_1\rangle = \frac{1}{\sqrt{d}} \sum_{x=0}^{d-1} |x\rangle \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right) \ . \tag{12.21}$$

After the simultaneous function computation with kick back, this yields – in analogy to Eq. (12.14) –

$$|\psi_1\rangle \xrightarrow{U_f} |\psi_2\rangle = \frac{1}{\sqrt{d}} \sum_{x=0}^{d-1} (-1)^{f(x)} |x\rangle \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right) \ . \tag{12.22}$$

We add a further transformation of the first register with Hadamard gates. The effect of $H^{\otimes n}$ will not be calculated in detail at this point. The following considerations are sufficient for our purposes: the effect of an individual Hadamard gate is $|0\rangle \xrightarrow{H} \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ or $|1\rangle \xrightarrow{H} \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. The state $|x\rangle$ of the x register is written in "binary form" as in Eq. (12.2). With $H^{\otimes n}$, we obtain

$$|x\rangle \xrightarrow{H^{\otimes n}} \frac{1}{d}\left(|0\rangle|0\rangle\ldots|0\rangle + |R(x)\rangle\right), \quad |R(x)\rangle \neq |0\rangle|0\rangle\ldots|0\rangle \ . \tag{12.23}$$

As the state of both registers, we thus find by evaluating Eq. (12.22)

$$\begin{aligned} |\psi_2\rangle &\xrightarrow{H^{\otimes n}\otimes \mathbb{1}} |\psi_3\rangle \\ &= \frac{1}{d}\left(\sum_{x=0}^{d-1} (-1)^{f(x)} \left[|0\rangle|0\rangle\ldots|0\rangle + |R(x)\rangle\right]\right) \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right) \ . \end{aligned} \tag{12.24}$$

Finally, we perform a measurement on the first register in the computational basis. The probability of obtaining the measurement result $(0,0,\ldots,0)$ is

$$p(0,0,\ldots,0) = \frac{1}{d^2}\left|\sum_{x=0}^{d-1} (-1)^{f(x)}\right|^2 \ . \tag{12.25}$$

It follows from this that

$$\begin{aligned} p(0,0,\ldots,0) &= 1 \quad \Leftrightarrow \quad f \text{ constant} \\ p(0,0,\ldots,0) &= 0 \quad \Leftrightarrow \quad f \text{ balanced .} \end{aligned} \tag{12.26}$$

We have carried out only one measurement on the first register. If the result $0,0,\ldots,0$ is found, f is constant. If any other measurement result occurs, f cannot be constant, i. e. f must be balanced.

For arbitrary register lengths n, it is thus sufficient in a quantum network to make only one query of the oracle. In a classical network, one would have to call up in sequence all the values of $f(x)$ for the N possible values of x. As soon as one obtains different results for two different x values, the function cannot be constant and is therefore balanced. In order to know with certainty that $f(x)$ is constant, the same result must be obtained in more than half of all cases, that is in at least $2^{n-1}+1$ cases. The number of queries necessary in the classical case thus increases exponentially with n.

12.5 Grover's Search Algorithm

The phone book problem The *phone book problem* involves finding the correct name in a telephone book for a given telephone number (e. g. 7581) out of all together d names.

x	Name	Telephone number
1	Jeeves	4892
2	Jones	1739
...	...	...
l	Smith	7581
...	...	...

The distribution of the telephone numbers is supposed to be random. The phone book is stored in the oracle. For the classical algorithm, the query is, "Does Jeeves have the number 7581?". The oracle answers in this case with "No". The names are queried in this manner one after another until "Smith" has been located. The use of the quantum parallelism in *Grover's algorithm* [Gro 96] again makes it possible to pose all the queries at once, so to speak. The readout is performed in this case by making use of amplitude amplification (see Sect. 12.3).

The telephone number 7581 corresponds to a function $f(x), x = 0,1,\ldots,d-1$ with the values

$$f(x) = 0 \text{ for } x \neq l \quad , \qquad f(x) = 1 \text{ for } x = l\,. \tag{12.27}$$

The quantum oracle allows the computation of $f : \{0,1\}^n \to \{0,1\}$ with $d = 2^n$ states $|x\rangle$ in the first register. The second register again consists of only one qubit. We let the state $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ be input to the second register and make use of the kick back as in Sect. 12.2. The state $\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ is the output. In the following, we list only the transformations in the first register. Then the effects of the quantum-mechanical function computation are described as in Eq. (12.13) by the unitary operator U_l.

$$|x\rangle \xrightarrow{U_l} (-1)^{f(x)}|x\rangle \tag{12.28}$$

U_l flips the state $|l\rangle$ in $-|l\rangle$ and leaves all the other states unchanged. One can therefore write this operator also in the form

$$U_l = \mathbb{1} - 2|l\rangle\langle l| \; . \tag{12.29}$$

The oracle is complemented by the action of further unitary operators. The unitary operator

$$U_R := 2|\Omega\rangle\langle\Omega| - \mathbb{1} \tag{12.30}$$

produces a "reflection" of $|\Omega\rangle$ of Eq. (12.15). It maintains the uniform superposition $|\Omega\rangle$ as in Eq. (12.15) and flips the sign of every vector orthogonal to $|\Omega\rangle$. Using Eq. (12.15), U_R can be written with Hadamard transformations in the form

$$U_R = H^{\otimes n}(2|0\rangle\langle 0| - \mathbb{1})H^{\otimes n} \tag{12.31}$$

with $|0\rangle = |0, 0, \ldots, 0\rangle$. The middle operator of the three produces a "reflection" of $|0\rangle$. It can be implemented using an n-bit Toffoli gate, which in turn can be constructed from 3-bit Toffoli gates. U_R can thus be implemented in terms of simple gates. More details are given in Problem 12.3.

We wish to explain the action of "reflection" of a state

$$|\phi\rangle = \sum_x a_x|x\rangle, \quad a_x \in \mathbb{R} \; . \tag{12.32}$$

The projection of $|\phi\rangle$ onto $|\Omega\rangle$ leads to

$$\langle\Omega|\phi\rangle = \frac{1}{\sqrt{d}} \sum_x a_x = \sqrt{d}\,\bar{a} \tag{12.33}$$

with the mean value of the amplitudes

$$\bar{a} := \frac{1}{d} \sum_x a_x \; . \tag{12.34}$$

This leads via the application of U_R to $|\phi\rangle$ with Eqs. (12.30) and (12.32) to the result

$$U_R|\phi\rangle = (2|\Omega\rangle\langle\Omega| - \mathbb{1})|\phi\rangle = 2\left(\sum_x |x\rangle\right)\bar{a} \; - \sum_x a_x|x\rangle = \sum_x (2\bar{a} - a_x)|x\rangle \; . \tag{12.35}$$

The amplitudes a_x of $|\phi\rangle$ are transformed according to $a_x \to 2\bar{a} - a_x$. This corresponds to a reflection of a_x at the mean value $\bar{a}$.

The sequence of the algorithm thus consists of the following: in a first iteration step, $U_R U_l$ acts upon $|\Omega\rangle$

$$|\phi_1\rangle = U_R U_l|\Omega\rangle \; . \tag{12.36}$$

In Fig. 12.8, the real amplitudes a_x are plotted against x. For $|\Omega\rangle$, they all have the value $\frac{1}{\sqrt{d}}$. Application of U_l flips a_l into $-a_l$ and leaves the remaining amplitudes unchanged.

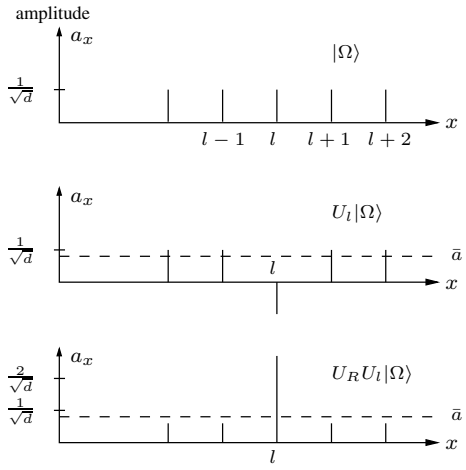


Figure 12.8: Amplitude amplification with Grover's algorithm (to be read from above).

The mean value $\bar{a}$ is shifted downwards in this process. The following transformation U_R reflects the values a_x about the new mean value and produces an amplification of the amplitude a_l in the state $|\Omega_1\rangle$. This is the result of the first cycle. Then $U_R U_l$ is once more applied to $|\phi_1\rangle$. The cycle is repeated several times. Finally, the x register is measured out bit by bit. Then the probability is greatest of finding the state $|l\rangle$ and hence the result l in dual notation. Grover's algorithm describes a situation in which the quantum computer yields the desired result not with certainty but only with a high probability. However, as a result of quantum parallelism, it is considerably faster than the repeated querying of a classical oracle,. A unitary transformation is a rotation in a complex space. Repeated applications can rotate a state increasingly near to a desired state. It can however also happen that the rotation goes beyond the desired state and on further repetition moves further and further away from it. It is therefore important to know for the application of Grover's algorithm when the iteration process must be discontinued (see Sect. 12.9).

A systematic classical search of the data bank requires a number of queries of the order of 2^n. This number increases exponentially with increasing n. In order to find the correct entry by means of Grover's algorithm with a high probability, $\sqrt{2^n}$ queries suffice (compare Sect. 12.9).

12.6 Shor's Factorisation Algorithm

Modern communication in the military and non-military area is based to an increasing extent on secure cryptography for the public transmission of keys and signatures. Up to the present time, the most important coding methods are based on the assumption that there is no effective factorisation of large numbers into prime numbers. The quantum algorithm of Shor ([Sho 94] and [Sho 97]) permits a very rapid factorisation in comparison to classical methods.

If it were technically possible to implement an effective quantum processor for this algorithm, this would have a serious effect on the security of secret data transmission and storage. This is one of the main reasons for the rapidly growing interest in quantum algorithms and in the construction of quantum computers. We want to introduce the quantum-mechanical factorisation algorithm here. It consists of a classical algorithm which contains stochastic elements (random elements), and the actual quantum algorithm for finding the period of a function. We begin with the classical part.

12.6.1 Reduction of Factorisation to the Search for a Period

The greatest common devisor as an ancillary quantity The fundamental theorem of arithmetic states that for every natural number $N > 1$, there is a unique *prime factorisation*

$$N = p_1^{n_1} p_2^{n_2} \dots p_k^{n_k} \tag{12.37}$$

with various distinct prime numbers $p_1, \dots, p_n$ and non-negative natural numbers $n_1, \dots, n_k$. Our goal is to find a fast algorithm for prime factorisation. We make use of the fact that there are already efficient methods for determining the *greatest common divisor* gcd(a, b) of two natural numbers a and b (cf. Sect. 12.9). The gcd(a, b) is the largest integer which is a divisor of both a and b.

$$\begin{aligned} 12: &\quad \text{divisor } 1, 2, 3, 4, 6, 12 \ . \\ 18: &\quad \text{divisor } 1, 2, 3, 6, 9, 18 \ . \\ \Rightarrow &\quad \gcd(12, 18) = 6 \ . \end{aligned} \tag{12.38}$$

Let $N \in \mathbb{N}$ be odd and not a prime number. For the prime factorisation of N, we search for a non-trivial divisor of N and apply the procedure successively to the factors of N thus obtained. For this purpose it suffices to find a natural number $b \neq 1$ which has at least one divisor in common with N; then we have with gcd(b, N) in particular also found a divisor of N. Such a situation occurs when there are natural numbers $c > N$ and $d > N$ which are *not divisible by* N, so that the equation

$$\frac{cd}{N} = m \tag{12.39}$$

is obeyed for a natural number m. Then it must be possible to cancel all the factors of the prime factorisation of N by some or all the factors of c and d. There thus exist a gcd(c, N) and a gcd(d, N). We compute both with the gcd algorithm and hence find factors of N at the same time.

The role of period determination How are we to find a relation of the type (12.39) for a given N? We first engage in some preliminary considerations. Let a be a natural number with $2 \leq a \leq N - 1$. We assume that

$$\gcd(a, N) = 1 \tag{12.40}$$

is obeyed – otherwise a divisor of N would already be known – and consider the function

$$f(x) := a^x \bmod N \ . \tag{12.41}$$

It can be shown (cf. Sect. 12.9) that this function $f(x)$ has a period r. This means that r is the smallest natural number for which

$$f(x+r) = f(x) \bmod N \tag{12.42}$$

holds. The period r depends on a. From

$$a^{x+r} = a^x a^r = a^x \bmod N\ , \tag{12.43}$$

it follows that

$$a^r = 1 \bmod N\ . \tag{12.44}$$

We give a simple example and compute in a modular fashion

$$(x, f(x) = 2^x \mod 3): \quad (1,2),\ (2,1),\ (3,2),\ (4,1), \ldots \tag{12.45}$$

The period is $r = 2$ and we have

$$a^r = 2^2 = 1 \bmod 3\ . \tag{12.46}$$

Let us assume that for a given a, the period r has already been found by a suitable procedure. And let us furthermore assume that the following two conditions are fulfilled:

$$r \text{ is even and} \tag{12.47}$$

$$a^{\frac{r}{2}} + 1 \neq 0 \bmod N\ . \tag{12.48}$$

Then owing to the condition (12.47), we can rearrange Eq. (12.44)

$$a^r - 1 = (a^{\frac{r}{2}} + 1)(a^{\frac{r}{2}} - 1) = 0 \bmod N\ . \tag{12.49}$$

The left-hand side of Eq. (12.49) must be a multiple of N. There is thus a natural number $m > 0$ such that we can write

$$\frac{(a^{\frac{r}{2}} + 1)(a^{\frac{r}{2}} - 1)}{N} = m\ . \tag{12.50}$$

From the requirement (12.48), it follows that $a^{\frac{r}{2}} + 1$ is not a multiple of N. Since r is the smallest number with the property (12.42), it must furthermore hold that

$$a^{\frac{r}{2}} - 1 \neq 0 \bmod N\ , \tag{12.51}$$

since otherwise due to Eq. (12.44), $\frac{r}{2}$ would already be the period. Then $a^{\frac{r}{2}} - 1$ is also not a multiple of N. On the other hand, Eq. (12.50) means that we can cancel all the factors of N on the left-hand side by factors in the numerator. This cancellation must occur in $(a^{\frac{r}{2}} + 1)$ and in $(a^{\frac{r}{2}} - 1)$, since otherwise one of these terms would be a multiple of N, in contradiction

to Eqs. (12.48) and (12.51). Both terms thus have common divisors with N. Among these is the largest, with

$$\gcd(a^{\frac{r}{2}}+1, N) \neq \binom{N}{1}, \quad \gcd(a^{\frac{r}{2}}-1, N) \neq \binom{N}{1} \ . \tag{12.52}$$

This statement of existence is our result. If we finally apply the algorithm for finding the gcd, we have obtained one or two factors of N.

We will thus be successful if we can, for a given N, find an a such that the two conditions (12.47) and (12.48) are fulfilled. The associated search is carried out by inputting random values for a and applying the algorithm until a value of a is found which fulfills the conditions. We are thus dealing with a *randomised algorithm* The factors of N obtained in this way are treated by the same procedure until finally the decomposition into prime factors is complete.

Flow diagram The scheme described above for the factorisation algorithm for the number N is shown as a flow diagram in Fig. 12.9. Only the search for the period, shown within a double frame, employs a quantum algorithm.

An example for N=15 We wish to factorise the number $N = 15$. The natural number a must therefore lie within the interval $2 \leq a \leq 14$. Since Eq. (12.40) must be obeyed, only the following choices for a are possible:

$$a \in \{2, 4, 7, 8, 11, 13, 14\} \ . \tag{12.53}$$

We choose randomly $a = 11$ and search for the period r of $f(x)$ as in Eq. (12.41).

$$\begin{aligned}
x = 0 \quad &: \quad 11^0 = 1 \bmod 15 \\
x = 1 \quad &: \quad 11^1 = 11 \bmod 15 \\
x = 2 \quad &: \quad 11^2 = 121 = 8 \cdot 15 + 1 = 1 \bmod 15 \\
x = 3 \quad &: \quad 11^3 = 1331 = 88 \cdot 15 + 11 = 11 \bmod 15 \ .
\end{aligned} \tag{12.54}$$

We find the period $r = 2$ and obtain with Eq. (12.52)

$$\gcd(11 + 1,\, 15) = 3, \qquad \gcd(11 - 1,\, 15) = 5 \ . \tag{12.55}$$

The numbers 3 and 5 are divisors of 15. The prime-number factorisation of 15 is thus $15 = 3 \cdot 5$. We were lucky with the choice $a = 11$.

If the random choice of a had instead been $a = 14$, we would obtain:

$$\begin{aligned}
x = 0 \quad &: \quad 14^0 = 1 \bmod 15 \\
x = 1 \quad &: \quad 14^1 = 14 \bmod 15 \\
x = 2 \quad &: \quad 14^2 = 196 = 13 \cdot 15 + 1 = 1 \bmod 15 \\
x = 3 \quad &: \quad 14^3 = 2744 = 182 \cdot 15 + 14 = 14 \bmod 15 \ .
\end{aligned} \tag{12.56}$$

The period is again $r = 2$ and therefore even. We take

$$\gcd(14 + 1, 15) = 15, \qquad \gcd(14 - 1, 15) = 1 \ . \tag{12.57}$$

This, however, violates the condition (12.52). In Fig. 12.9 (flow diagram), the flow cycle leads back to the beginning. A new value for a must be chosen.

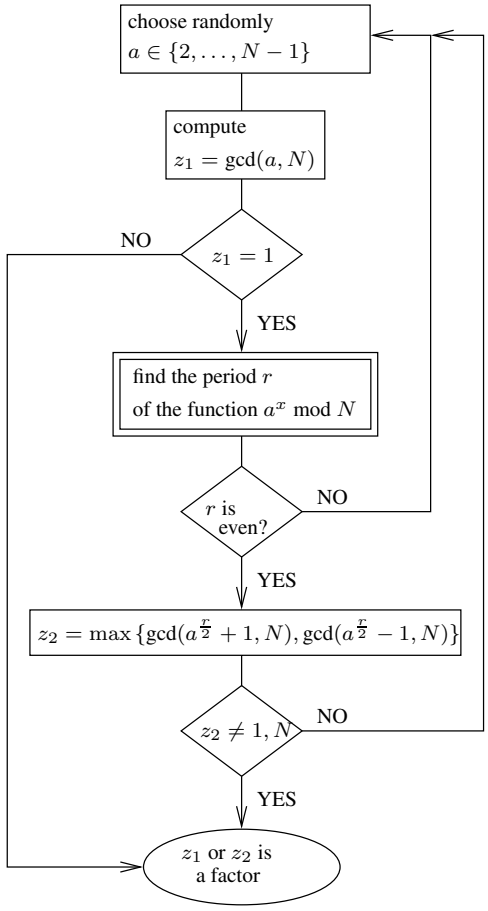


Figure 12.9: A flow diagram of the factorisation algorithm for the number N. Only the part within the double frame is carried out using the quantum computer.

12.6.2 The Quantum Algorithm for Determining the Period

Our remaining task is to determine the period of the function $f(x) = a^x \bmod N$. We again use two registers of lengths n and m. In the first register, the $d = 2^n$ basis vectors $|x\rangle$ of $\mathcal{H}_2^{\otimes n}$ as well as every superposition $|\phi\rangle \in \mathcal{H}_2^{\otimes n}$ can be input. In the second register, $f(x) \bmod N$ is stored. The length m must be chosen in such a way that the dimension is $2^m \geq N$. States in this register are denoted as $|\chi\rangle \in \mathcal{H}_2^{\otimes m}$. The composite state $|\psi\rangle \in \mathcal{H}_2^{\otimes n} \otimes \mathcal{H}_2^{\otimes m}$ of all the registers is in general entangled.

1st step: initialisation In the beginning, both registers are in the respective basis states $|0\rangle$ (see Fig. 12.10). As starting point for the utilisation of quantum parallelism, in a first step the state $|\psi\rangle$ is brought in the well-known manner into the uniform superposition of the basis

states of $\mathcal{H}_2^{\otimes n}$

$$|\psi_1\rangle = |\Omega\rangle|0\rangle = \frac{1}{\sqrt{d}} \sum_{x=0}^{d-1} |x\rangle|0\rangle \ . \tag{12.58}$$

Example: N=15
$1st$ register: $n = 3$ qubits for the numbers 0 to 7. Thus, $d = 8$.
$2nd$ register: $m = 4$ qubits for the numbers 0 to $N = 15$.

$$|\psi_1\rangle = \frac{1}{\sqrt{8}} \left(|0\rangle + |1\rangle + \ldots |7\rangle\right) |0\rangle \, . \tag{12.59}$$

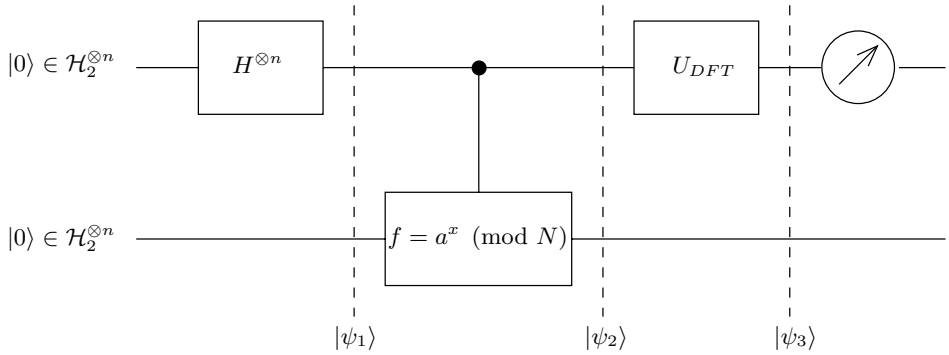


Figure 12.10: The quantum circuit for determining the period.

2nd step: computation of $f(x)$ in the second register The result of the unitary transformation U_f of Eq. (12.41) for the computation of $f(x)$ is the entangled state

$$|\psi_2\rangle = \frac{1}{\sqrt{d}} \sum_{x=0}^{d-1} |x\rangle|a^x \mathrm{mod}\ N\rangle \ . \tag{12.60}$$

Example N = 15 and a = 11:
The random choice of a is supposed to have led to 11. Then for $|\psi_2\rangle$, we find

$$\begin{aligned} |\psi_2\rangle &= \frac{1}{\sqrt{8}}(|0\rangle|1\rangle + |1\rangle|11\rangle + |2\rangle|1\rangle + |3\rangle|11\rangle \\ &\quad + \cdots + |7\rangle|11\rangle) \end{aligned} \tag{12.61}$$

$$\begin{aligned} |\psi_2\rangle &= \frac{1}{\sqrt{8}}\{(|0\rangle + |2\rangle + |4\rangle + |6\rangle)\,|1\rangle \\ &\quad + (|1\rangle + |3\rangle + |5\rangle + |7\rangle)\,|11\rangle\} \ . \end{aligned} \tag{12.62}$$

The period we are seeking can be read off the sequence of x values for which the resulting functionals repeat. Eq. (12.62) is ordered according to the function values in the second register. The period (here $r = 2$) is stored in the different states of the first register, which are obtained in the summary (12.62). On expanding these states in terms of basis states, the lowest numbers which occur can exhibit shifts of $l \in \mathbb{N}$ relative to 0, which are also called *offsets*. In Eq. (12.62), we have for the numbers of the basis vectors $(0, 2, 4, 6)$ and $(1, 3, 5, 7)$. The corresponding offset is $l = 0$ or $l = 1$.

We can rearrange the right-hand side of Eq. (12.60) and sort as in Eq. (12.62) according to the offset l:

$$|\psi_2\rangle \sim \sum_l |\phi_2(l)\rangle |a^l \bmod N\rangle \tag{12.63}$$

with

$$|\phi_2(l)\rangle = \left[\frac{d}{r}\right]^{-\frac{1}{2}} \sum_{j=0}^{[\frac{d}{r}]-1} |l + jr\rangle \; . \tag{12.64}$$

$[\frac{d}{r}]$ is here the largest natural number which is smaller than or equal to $\frac{d}{r}$. In the example, $[\frac{d}{r}] = 4$.

Equations (12.62) and (12.64) show that the period r is already stored in the states of the first register after the $2nd$ step. We carry out a measurement on the second register and select according to the results of the measurement. We then perform a measurement on the first register. From the corresponding results, the period can be read off. In our example, after the selection of $l = 0$ (i. e. the measured result 1 in the second register) a state is present which leads to the results $0, 2, 4, 6$ in the first register. The period $r = 2$ is clear. An analogous result is obtained for $l = 1$ (i. e. the measured result 11 in the second register). This procedure however becomes more and more laborious with increasing N, since then more and more different measured values are found in the measurement on the second register. They must all occur frequently in order that many systems are present after the selection of a measured value and the measurements on the first register hence permit a conclusion to be drawn with respect to the period r. We therefore make use of a different procedure. It is based on the fact that via a unitary transformation, all the offsets can be uniformly set to 0. It then suffices to carry out measurements only on the first register in order to be able to determine the period.

3rd step: quantum Fourier transformation The procedure consists of making use of a Fourier transform of the first register (see Fig. 12.10) to convert the offset l into a phase factor which is irrelevant for the quantum-mechanical measurement. The unitary operator for *quantum Fourier transforms* acts as follows:

$$U_{QFT}|x\rangle = \frac{1}{\sqrt{d}} \sum_{z=0}^{d-1} \exp\left(2\pi i \frac{xz}{d}\right) |z\rangle \, , \quad z \in \mathbb{N} \; . \tag{12.65}$$

We apply U_{QFT} to the state $|\phi_2(l)\rangle$ from Eq. (12.64) and discuss first the *special case*

$$\frac{d}{r} \in \mathbb{N} \, . \tag{12.66}$$

Then we have

$$|\phi_2(l)\rangle = \sqrt{\frac{r}{d}} \sum_{j=0}^{\frac{d}{r}-1} |l + jr\rangle \tag{12.67}$$

with

$$U_{QFT}|\phi_2(l)\rangle = \sum_{z=0}^{d-1} \tilde{f}(z)|z\rangle \tag{12.68}$$

and

$$\begin{aligned} \tilde{f}(z) &= \frac{\sqrt{r}}{d} \sum_{j=0}^{\frac{d}{r}-1} \exp\left(2\pi i \frac{(l+jr)z}{d}\right) \\ &= \frac{\sqrt{r}}{d} \left[\sum_{j=0}^{\frac{d}{r}-1} \exp\left(2\pi i j \frac{rz}{d}\right)\right] \exp\left(2\pi i \frac{lz}{d}\right) . \end{aligned} \tag{12.69}$$

We investigate the factor $[\dots]$ further. It represents a geometric series.

$$[\dots] = \frac{\exp\left(2\pi i \frac{rz}{d} \cdot \frac{d}{r}\right) - 1}{\exp\left(2\pi i \frac{rz}{d}\right) - 1} = \frac{\exp(2\pi i z) - 1}{\exp(2\pi i \frac{r}{d} z) - 1} . \tag{12.70}$$

The series has the value 1 when the natural number z is a multiple of the natural number $\frac{d}{r}$, and otherwise it is zero. $\tilde{f}(z)$ is non-vanising only at the points $z = k\frac{d}{r}$ with $k \in \mathbb{N}$.

The new interval of values for z is transferred to the states $|z\rangle$. This leads to

$$U_{QFT}|\phi_2(l)\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} \exp\left(2\pi i \frac{lk}{r}\right) |k\frac{d}{r}\rangle . \tag{12.71}$$

The offset l is now located within the phase. The period r is in the enumeration of the states ($k\frac{d}{r} \in \mathbb{N}$). The composite state is, up to a normalisation factor, given by

$$|\psi_3(l)\rangle \sim \sum_l U_{QFT}|\phi_2(l)\rangle \otimes |a^l \bmod N\rangle . \tag{12.72}$$

4th step: measurements on the first register We measure on the first register (compare Fig. 12.10) in the computational basis and repeat the whole procedure several times. According to Eq. (12.71), only the results $x_k = k\frac{d}{r}$ with $k = 0 \dots, r-1$ occur with non-vanishing constant probabilities (see Fig. 12.11a). From this, we can read off $\frac{d}{r}$. Since d is known, the period r is determined and the procedure terminates.

Example $\mathbf{N = 15}$ ***and*** $\mathbf{a = 11}$:

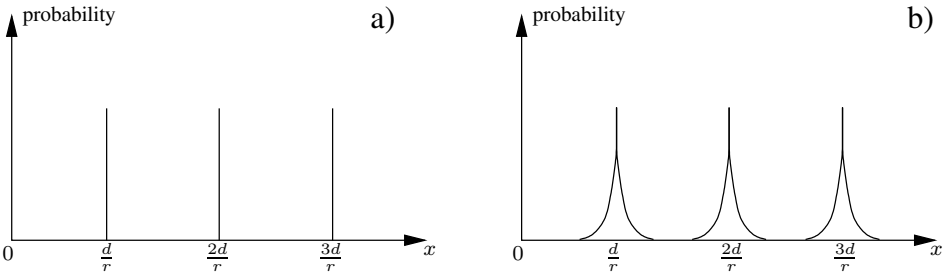


Figure 12.11: The probability distribution of the measured results x in the first register for the Shor algorithm.

We have $d = 8$.
The summands in Eq. (12.62) have the offsets $l = 0$ *and* $l = 1$. *The unitary transformation* U_{QFT} *leads to*

$$|\psi_3\rangle = \frac{1}{\sqrt{4}} \left\{ |0\rangle(|1\rangle + |11\rangle) + |4\rangle(|1\rangle + e^{i\pi}|11\rangle) \right\} \ . \tag{12.73}$$

With the probability $\frac{1}{2}$, *the result* 0 *will be measured on the first register. This yields no information. If the value* 4 *belonging to* $k = 1$ *is measured, it follows from* $4 = \frac{d}{r}$ *with* $d = 8$ *that the period is* $r = 2$. *This completes the determination of the period.*

We add that with only one measurement with the measured value x', there is a certain probability that a k' is found that has no common divisor with r, $(\gcd(k', r) = 1)$

$$\frac{x'}{d} = \frac{k'}{r} \ . \tag{12.74}$$

Then one cancels in the quotient x'/d until an irreducible fraction is obtained (there are fast algorithms to accomplish this) and reads off r. With a certain probability, a single measurement thus suffices for the determination of the period. If the measurement does not lead to a k with the property described above, then the procedure must be repeated. It can be shown that the computer time required for this procedure is less than that for classical methods of determining the period.

The general case If r is not a divisor of d and hence the special case (12.66) is not at hand, we nevertheless expect that the probability distribution is concentrated around values of x which are "nearly" a multiple of $\frac{d}{r}$ (cf. Fig. 12.11b). This can in fact be proved. A suitable evaluation procedure can be given (see Sect. 12.9). The time required to carry out the quantum Fourier transform in a d-dimensional space is of the order of $(\log d)^2$. The classical fast Fourier transform requires of the order of $d \log d$. It is thus inferior (compare Sect. 12.9).

12.7 Quantum Error Correction Using Non-local Measurements

As in classical computers, real quantum computers are subject to errors. We shall see in detail in Chap. 15 how the interactions with the environment lead to decoherence. The pure states on which the computations are based are transformed into mixtures. A different disturbance occurs because the quantum gates of which the computer is constructed do not perform perfectly and possibly carry out perturbed unitary transformations. Especially drastic disturbances occur when states in register positions switch (e. g. $|0\rangle \rightarrow |1\rangle$) or change in phase (e. g. $|0\rangle \rightarrow -|0\rangle$). Both in the computer as well as during data transfer in channels, quantum information must be protected against losses by detection of errors (diagnosis) and their correction (therapy).

The usual procedures for error correction of classically-processed information cannot be transferred to quantum computing, since they are based on the copying of states and on local measurements. There are however no universal copiers for quantum states, and local measurements destroy entanglement. Quantum procedures for error correction are required. We describe some of them.

The *quantum error-correcting codes* (QECC) are an example of the use of entanglement and of non-local measurements (cf. Sect. 9.2) as ancillary procedures. The fundamental idea consists of storing information in a redundant manner. Entangled states are formed. In the simplest case, single errors occur locally, i. e. in subsystems and individual registers. Then these errors can be detected by non-local measurements and the locally-hidden information can again be restored. We will demonstrate this with some examples.

12.7.1 Bit Flip Errors

A *bit-flip error* $|0\rangle \rightarrow |1\rangle, |1\rangle \rightarrow |0\rangle$ occurs locally at a single register position. The state remains a pure state. One can protect against this type of error by redundantly coding a single qubit by means of three qubits in the following way:

$$|0\rangle \rightarrow |\bar{0}\rangle := |0,0,0\rangle \qquad |1\rangle \rightarrow |\bar{1}\rangle := |1,1,1\rangle \; . \tag{12.75}$$

Thus, the state $|\varphi\rangle$ is converted into the entangled state $|\phi\rangle$

$$|\varphi\rangle := c_0|0\rangle + c_1|1\rangle \quad \rightarrow \quad |\phi\rangle := c_0|0,0,0\rangle + c_1|1,1,1\rangle \; . \tag{12.76}$$

The corresponding quantum circuit can be constructed using two CNOT gates as in Fig. 12.12.

We will no longer denote the product Hilbert spaces by capital letters A, B etc. , but instead we number them serially: $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)} \otimes \ldots$. The states $|\bar{0}\rangle$ and $|\bar{1}\rangle$ are both eigenstates of $\sigma_z^{(1)}\sigma_z^{(2)}$ and $\sigma_z^{(2)}\sigma_z^{(3)}$ with the eigenvalue $+1$. The corresponding non-local measurements in the computational basis yield the measured value $+1$ (see Sect. 9.2).

A bit flip in a register leads for example to the entangled state

$$|\phi'\rangle := c_0|1,0,0\rangle + c_1|0,1,1\rangle \; . \tag{12.77}$$

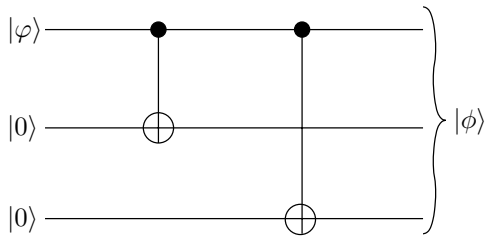


Figure 12.12: Generating a redundant coding.

A non-local measurement of the observable $\sigma_z^{(1)}\sigma_z^{(2)}$ now yields the value -1, while that of the observable $\sigma_z^{(2)}\sigma_z^{(3)}$ yields $+1$. This permits the unambiguous conclusion that a bit flip has occurred in the first register. It is important that in the non-local measurement, the state $|\phi'\rangle$ was not modified. Hence, a unitary transformation $\sigma_x^{(1)} \otimes \mathbb{1}^{(2)} \otimes \mathbb{1}^{(3)}$, which produces a bit flip in the state of the first register, leads back to the primary state $|\phi\rangle$. The pairs of values which are obtained from the measurement of $\sigma_z^{(1)}\sigma_z^{(2)}$ and $\sigma_z^{(2)}\sigma_z^{(3)}$ are: $(+1,+1), (-1,+1), (-1,-1), (+1,-1)$. They correspond in order to: no flip, a flip in the first, in the second, or in the third register digit. The bit flip is reversed by another bit flip at the same position.

Unitary transformations can be close to the identity operation. We consider the example ($|\epsilon| \ll 1$)

$$\begin{aligned} |\phi\rangle \rightarrow |\phi''\rangle \quad &= \quad c_0' \left(|0,0,0\rangle + \epsilon|1,0,0\rangle\right) \\ &\quad + c_1' \left(|1,1,1\rangle + \epsilon|0,1,1\rangle\right) \ . \end{aligned} \tag{12.78}$$

By measurement of $\sigma_z^{(1)}\sigma_z^{(2)}$ and $\sigma_z^{(2)}\sigma_z^{(3)}$, the result $(+1,+1)$ will be obtained with the probability $1 - 2|\epsilon|^2$ and $|\phi''\rangle$ is projected back onto the state $|\phi\rangle$. The result $(-1,+1)$ occurs with the probability $2|\epsilon|^2$ and the measurement leads to $|\phi'\rangle$. The result of the measurement indicates the final state and allows correction of the error where necessary.

12.7.2 Phase Flip Errors

A qubit is coded with 9 qubits, which are combined into clusters of 3 qubits:

$$\begin{aligned} |0\rangle \quad &\rightarrow \quad |\bar{0}\rangle = \frac{1}{2^{3/2}} \left(|0,0,0\rangle + |1,1,1\rangle\right) \left(|0,0,0\rangle + |1,1,1\rangle\right) \left(|0,0,0\rangle + |1,1,1\rangle\right) \\ |1\rangle \quad &\rightarrow \quad |\bar{1}\rangle = \frac{1}{2^{3/2}} \left(|0,0,0\rangle - |1,1,1\rangle\right) \left(|0,0,0\rangle - |1,1,1\rangle\right) \left(|0,0,0\rangle - |1,1,1\rangle\right) \ . \end{aligned} \tag{12.79}$$

Every cluster has a redundant bit coding. A single bit flip can be detected as in Sect. 12.7.1 and eliminated. A *phase flip error* which gives rise to a sign change in one of the 9 register positions can be determined in the following way: one measures the two non-local 6-qubit

observables

$$\begin{aligned}&\sigma_x^{(1)}\sigma_x^{(2)}\sigma_x^{(3)}\sigma_x^{(4)}\sigma_x^{(5)}\sigma_x^{(6)}\\&\sigma_x^{(4)}\sigma_x^{(5)}\sigma_x^{(6)}\sigma_x^{(7)}\sigma_x^{(8)}\sigma_x^{(9)}\end{aligned} \tag{12.80}$$

which are formed from the bit flip operator σ_x. The states $|\bar{0}\rangle$ and $|\bar{1}\rangle$ are eigenstates with the eigenvalue $+1$.

If a phase flip occurs in one of the registers, then the value resulting from a measurement of $\sigma_x\sigma_x\sigma_x$ for this cluster changes by a factor of -1. By measuring the operators (12.80), one can determine as for the bit flip error in which cluster the phase flip has occurred. The error correction then consists of carrying out a σ_z transformation within this cluster at some register position. This restores the initial state.

Errors at all the register positions We consider the possibility that an error can occur at all the register positions, but require that the resulting state be obtainable by a unitary transformation $U = \mathbb{1} + O(\epsilon)$ with $\epsilon \ll 1$. It has the quite general structure

$$U = \mathbb{1} + i\epsilon_x\sigma_x + i\epsilon_y\sigma_y + i\epsilon_z\sigma_z \ . \tag{12.81}$$

The individual terms cause bit flips, phase flips, or both together. We consider the coding (12.79). The diagnosis for phase flips and bit flips is again carried out and this will already make it highly probable that the state is projected back onto the unperturbed original state. With a smaller probability $|\epsilon|^2$, a bit flip or a phase flip will have occurred in a register. This error will be recognised and corrected in the known manner. In the case of errors at two or more positions, this is however not possible. This situation, however, occurs only with a probability of $|\epsilon|^4$ or less.

12.8 The Components of the Quantum Computer*

We limit ourselves in this section to giving only the relevant results. For proofs, we refer to the literature.

Universal gates Quantum gates carry out unitary transformations. They are wired together into circuits via quantum wires (undisturbed propagation). A finite number of quantum gates is called a *universal quantum gate* when every unitary transformation can be performed with it on every arbitrary qubit register. Unitary transformations merge into a continuum. Universality thus requires that every unitary transformation can be approximated with arbitrary precision. Universal gates are the components from which a quantum computer can be fabricated. They permit any quantum algorithm to be computed. Attempts to construct them are being made using *ion traps*, *cavity QED*, *quantum dots*, and *trapped atoms* (see [EMY 02], also for additional literature).

*The sections marked with an asterisk * can be skipped over in a first reading.

The following results are important in this connection:

- Every unitary operator on a multidimensional space can be described exactly as a product of gates which act upon only two qubits (2-qubit gates ([DiV 95]). To implement such gates, the interaction between only two physical systems suffices.
- The CNOT gate, together with simple single-qubit gates such as phase gates and rotations, form a universal set of gates ([BBC 95]).

$\sqrt{\text{SWAP}}$ gates The $\sqrt{\text{SWAP}}$ gate, complemented by single-qubit gates, is likewise universal and plays a similarly important role in the attempts to experimentally implement quantum computers as does the CNOT gate (see [LD 98] and [EMY 02]). The unitary $\sqrt{\text{SWAP}}$ gate is defined by its action on the computational basis:

$$|0^A, 0^B\rangle \xrightarrow{\sqrt{\text{SWAP}}} |0^A, 0^B\rangle \tag{12.82}$$

$$|0^A, 1^B\rangle \xrightarrow{\sqrt{\text{SWAP}}} \frac{1+i}{2}|0^A, 1^B\rangle + \frac{1-i}{2}|1^A, 0^B\rangle \tag{12.83}$$

$$|1^A, 0^B\rangle \xrightarrow{\sqrt{\text{SWAP}}} \frac{1-i}{2}|0^A, 1^B\rangle + \frac{1+i}{2}|1^A, 0^B\rangle \tag{12.84}$$

$$|1^A, 1^B\rangle \xrightarrow{\sqrt{\text{SWAP}}} |1^A, 1^B\rangle\,. \tag{12.85}$$

This corresponds to the matrix representation

$$\sqrt{\text{SWAP}} \leftrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1+i}{2} & \frac{1-i}{2} & 0 \\ 0 & \frac{1-i}{2} & \frac{1+i}{2} & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \tag{12.86}$$

in the computational basis. The relation $\sqrt{\text{SWAP}}\sqrt{\text{SWAP}} = \text{SWAP}$ (cf. Sect. 7.8.2) justifies the name. The CNOT operator can be reduced to the $\sqrt{\text{SWAP}}$ operator by making use of simple single-qubit operators. Along with the Hadamard gate from Eq. (3.53), we need the operator

$$\hat{\sigma} := e^{-i\frac{\pi}{4}\mathbb{1}} e^{-i\frac{\pi}{2}\sigma_z} \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & -i \end{pmatrix} \tag{12.87}$$

(the matrix representation is found up to an unimportant phase factor). With $\hat{\sigma}$, we obtain

$$\begin{aligned} \hat{U}^{AB} &:= (\sigma^A)^{-1}\sigma^B\sqrt{\text{SWAP}}(\sigma^A)^2\sqrt{\text{SWAP}} \\ &\leftrightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \end{aligned} \tag{12.88}$$

The Hadamard gate then leads to CNOT

$$\text{CNOT} = H^A \hat{U}^{AB} H^A \tag{12.89}$$

as in Eq. (7.98). This demonstrates the universality of the $\sqrt{\text{SWAP}}$ gate.

12.9 Complementary Topics and Further Reading

- Review articles on quantum computing: [Ben 95], [Bar 96], [PVK 96], [Bar 98], [CEM 98], [DE 98], [Pre 98, Chap. 6], [Ste 98], [VP 98], [Bra 99a], [Joz 00], [NC 00, Part II], [RP 00], [EHI 01], [CB 02], [GM 02], [Lom 02a], [Wer 06].
- Quantum computing; the classics: [Deu 85], [Sho 94].
- Books about quantum computing: [BDM 98], [WC 98], [Bra 99], [Bro 99], [Gru 99], [Pit 00], [CP 01], [Hir 01], [DM 02], [KSV 02], [LB 02], [Lom 02], [SS 04], [BCS 04].
- Books with review articles on the theory of quantum computers: [LSP 98], [Bra 99], [CB 02], [GM 02], [Hei 02], [Lom 02].
- Books with overviews of the attempts to implement quantum computers experimentally: [SS 04].
- Books or review articles on the experimental implementation of quantum computers: [Pel 98], [Gru 99, Chap. 7.6], [BEZ 00], [CLK 00], [DiV 00], [NC 00, Part II], [DM 02], [SS 04].
- Books with review articles on the experimental implementation of quantum computers: [LSP 98], [Bra 99], [BEZ 00], [DM 02], [Hei 02].
- An introduction to the implementation of quantum information processing with ions in traps ([Bla 06]) and with photons ([Wei 06]).
- The number of queries required with Grover's algorithm: [BBH 98], [Pre 98, Chap. 6], [EHI 01].
- Description of other search algorithms: [Gru 99, Chap. 3].
- Review article on the Shor algorithm: [EJ 96]. For details of the calculation see [NC 00].
- The Euclidian algorithm for the determination of the greatest common divisor: [NC 00, Appendix 4].
- When the condition (12.40) is fulfilled, the function $f(x)$ of Eq. (12.42) has a period. A proof using the theorem of Euler and Fermat: [HW 79].
- The quantum Fourier transform if Eq. (12.66) is not obeyed: [EJ 96], [Pre 98, Chap. 6.9.1].
- Computing times for classical and quantum-mechanical Fourier transforms: [Pre 98, Chap. 6].
- The Toffoli gate is a universal reversible gate. Every unitary transformation can be carried out by a combination of Toffoli gates: [Pre 98, Chap. 6], [Gru 99, Chaps. 1.7.1 and 3.1], [Hir 01, Chap. 2.3.2].

- Knowledge about computational complexity makes possible the estimation of limits and advantages of quantum computers: [Mer 02].
- Error correcting quantum codes : [Pre 98, Chap. 7], [Pre 98 a], [Gru 99, Chap. 7], [Pre 99].
- Decoherence via coupling to the environment destroys the unitary evolution in quantum computers and is therefore an important source of errors: [Bar 96], [PSE 96].
- Decoherence-free subspaces and systems are considered to be one of the most promising solutions to the decoherence problem in quantum computing: [LW 03].
- One can implement a Deutsch gate with nearly every 2-qubit gate ([Bar 95]). Hence these 2-qubit gates are universal.
- We have described quantum computation as a unitary evolution followed by measurements with which classical information is read out. A different form of dynamics is measurement dynamics. There are procedures for quantum computation in which the individual computational steps consist of measurements. These *measurement-based models* have no classical analogue. For a review, see: [Joz 05].

12.10 Problems for Chapter 12

Prob. 12.1 [for 12.2]: Given: a unitary mapping $U_y : \mathcal{H}_2 \to \mathcal{H}_2$ which maps every state $|\psi\rangle$ onto $(-1)^y|\psi\rangle$ for a given, fixed y. Construct a quantum algorithm for finding y. Use two Hadamard gates and a controlled U_y gate.

Prob. 12.2 [for 12.4.1]: The parity par(f) of a function $f : \{1,2\} \to \{-1,+1\}$ is defined by par$(f) := f(1)f(2)$. Construct a quantum algorithm which permits the computation of par(f) by means of a single query of the oracle (i. e. by using the black box for the function f only once).

Prob. 12.3 [for 12.4.2]:

(i) Show that the effect of the Hadamard transformation of n qubits $H^{(n)} = H \otimes H \otimes \cdots \otimes H$ on $|x\rangle$ has the form

$$|x\rangle \xrightarrow{H^{(n)}} \frac{1}{\sqrt{d}} \sum_{y=0}^{d-1} (-1)^{(x \cdot y)} |y\rangle \; . \tag{12.90}$$

An equally-weighted superposition of the basis vectors of $\mathcal{H}_2^{(2)}$ with signs $+1$ and -1 results. Here, $(x \cdot y)$ is the "vectorial inner product" of the register states

$$(x \cdot y) = x_{n-1}y_{n-1} + x_{n-2}y_{n-2} + \ldots + x_0 y_0 \; . \tag{12.91}$$

(ii) Write $|\psi_3\rangle$ from Eq. (12.24) in complete form.

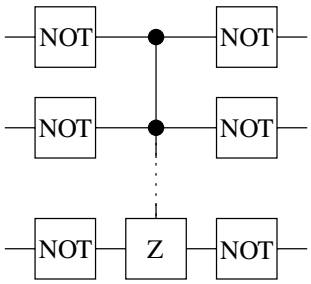


Figure 12.13: A quantum circuit for the "reflection" on $|0\rangle$.

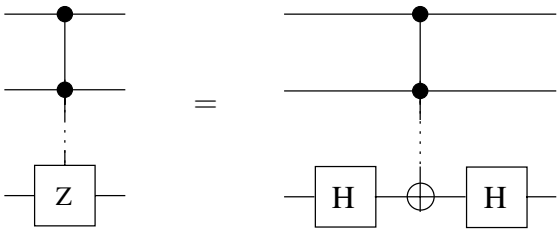


Figure 12.14: Reduction of the controlled Z gate to an n-bit Toffoli gate.

(iii) A black box computes the function $f_a : \{0,1\}^n \to \{0,1\}$ defined by

$$f_a(x) = (a \cdot x) \; . \tag{12.92}$$

Find a quantum algorithm which is able to distinguish the 2^n functions f_a by means of a single computation of f_a and a measurement of the x register (Bernstein-Vazirani problem). Start with the expression for $|\psi''\rangle$ obtained in part (ii) of this problem.

Prob. 12.4 [for 12.5]: For which value of N in Grover's algorithm is the state being sought, $|l\rangle$, already determined with certainty after one pass?

Prob. 12.5 [for 12.5]:
a) Show that the operator $2|0\rangle\langle 0| - \mathbb{1}$, which causes a reflection on $|0\rangle \in \mathcal{H}_2^{\otimes n}$, can be implemented up to a global minus sign by a quantum circuit which is based on a multiply-controlled Z gate (see Tab. 3.1 and Fig. 12.13).
b) The controlled Z gate can itself be constructed with an n-bit Toffoli gate (see Fig. 7.6) and Hadamard gates.

© 2007 WILEY-VCH Verlag GmbH & Co.

13 Generalised Measurements, POVM

We first give an example of the general dynamics of open systems and then go on to generalised measurements. In the next chapter, we will generalise the description of unitary evolution.

In quantum theory, not only projective measurements and unitary evolution are possible. Generalised measurements describe more complicated situations for measurements and open up new possibilities. Using as an example the non-optimal Stern-Gerlach experiment, we introduce such measurements, as well as the POVM measurements.

13.1 The Function of a Generalised Dynamics of Open Quantum Systems

13.1.1 Problems

Up to now, we have considered the extension of a quantum system S^A by a second system S^B to form a bipartite system S^{AB}. Interactions and entanglement between the two subsystems are permitted. This makes S^A into an *open system*. The composite system S^{AB} is however assumed once again to be a closed system. Like all closed systems, it can pass through a unitary dynamic evolution (unitary dynamics) and one can carry out projective measurements on it (measurement dynamics). The rules for both forms of dynamics were formulated in terms of the postulates in a previous chapter (cf. Sect. 7.3.1). In the coming chapters, we shall introduce a different point of view and place the openness of the system S^A at its centre. The dynamic evolution between preparation and measurement (compare Fig. 2.4) was termed the *transformation dynamics* in Chap. 2.

The following problems are to be dealt with in terms of a generalisation of unitary dynamics:

(i) Which form is assumed by the transformation dynamics of S^A in the particular situation that S^A is a subsystem of a closed system S^{AB} which itself is passing through a unitary evolution?

(ii) What structure does the most general physical transformation dynamics of an open system S^A have? This question is to be answered as far as possible without explicit reference to a second system (i. e. to S^B).

(iii) Can this – no longer necessarily unitary – generalised dynamics of S^A always be understood as the result of the unitary dynamics of a composite system S^{AB} which has been

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

extended to include an *ancilla* S^B? If this is possible, we would at the same time have discovered a procedure for the experimental implementation of the generalised dynamics of S^A.

Thus far, we have generalised the unitary dynamics. Analogously, one can generalise the projective measurement dynamics. We start again with the entangled composite system S^{AB}.

(i) A projective measurement on the second system S^B transforms a system S^A with which it is entangled into corresponding new states with associated probabilities.

(ii) If we disregard the possible existence of a second system S^B and consider only S^A, what is the most general physical structure of a measurement on a system S^A?

(iii) Can this – not necessarily projective – general measurement on S^A always be considered to be the result of a projective measurement on a system S^B which is entangled with S^A? A suitable entanglement with an ancilla S^B would then provide the basis for the implementation of a general measurement on S^A.

We will first use a simple example to read off which answers can be expected to the above two questions (i). We then turn to generalised measurements and the POV measures. In Chap. 14, we consider quantum operations as generalisations of the unitary dynamics and supplement this in Chap. 16 with proofs which were left off in earlier chapters.

13.1.2 A Simple Example

Generalised transformation dynamics In order to become familiar with the structures which we can expect, we first discuss a mathematically simple example from which one can already read off the general structure of the generalised dynamics and generalised measurements[1]. We begin with the unitary dynamics of a composite system S^{AB} which is composed of the subsystems S^A and S^B. We assume the initial state in $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$ to be the product state

$$\rho^{AB} = \rho^A \otimes |0^B\rangle\langle 0^B| \,. \tag{13.1}$$

It is transformed by the unitary transformation

$$U^{AB} = \frac{1}{\sqrt{2}} \left(\sigma_x^A \otimes \mathbb{1}^B + \sigma_y^A \otimes \sigma_x^B\right) \tag{13.2}$$

into the entangled composite state

$$\begin{aligned} \rho'^{AB} &= U^{AB} \rho^{AB} U^{AB\dagger} \\ &= \frac{1}{2}\Big\{\sigma_x^A \rho^A \sigma_x^A \otimes |0^B\rangle\langle 0^B| + \sigma_y^A \rho^A \sigma_y^A \otimes |1^B\rangle\langle 1^B| \\ &\quad + \sigma_x^A \rho^A \sigma_y^A \otimes |0^B\rangle\langle 1^B| + \sigma_y^A \rho^A \sigma_x^A \otimes |1^B\rangle\langle 0^B|\Big\} \,. \end{aligned} \tag{13.3}$$

The open system S^A is here transformed into the state

$$\rho^A \to \rho'^A = \mathrm{tr}_B[\rho'^{AB}] = \frac{1}{2}\left\{\sigma_x^A \rho^A \sigma_x^A + \sigma_y^A \rho^A \sigma_y^A\right\} \,. \tag{13.4}$$

[1] It could be helpful while reading this chapter to occasionally consult Table 14.1.

The unitary, entangled evolution of the composite system thus causes an evolution of the subsystem S^A which is non-unitary but which conserves the trace; it can be written using a superoperator $\mathcal{E}^A$ in the form of an *operator-sum decomposition* or *operator-sum representation*

$$\rho'^A = \mathcal{E}(\rho^A) = \sum_{i=1}^{2} K_i^A \rho^A K_i^{A\dagger} \; . \tag{13.5}$$

Here,

$$K_1^A := \frac{1}{\sqrt{2}}\sigma_x^A, \qquad K_2^A := \frac{1}{\sqrt{2}}\sigma_y^A \; . \tag{13.6}$$

The K_i^A are called *Kraus operators*, *operation elements* or *decomposition operators*. They obey the completeness relation

$$\sum_{i=1}^{2} K_i^{A\dagger} K_i^A = \mathbb{1} \; . \tag{13.7}$$

In anticipation of the proof which will be given later, we make the following statement: *The dynamics of the open system S^A, in general non-unitary, can be described by a superoperator $\mathcal{E}$, which acts only on the state ρ^A of S^A. For $\mathcal{E}$, there is an operator-sum representation.*

One can imagine the evolution (13.4) of the initial system S^A also to have been produced in a non-unitary fashion whereby the unitary operators σ_x^A or σ_y^A are applied to S^A with the probabilities $\frac{1}{2}$. This would be a very different process physically from that described by U^{AB} in Eq. (13.2). This non-unitary evolution of S^A can hence be implemented as the unitary evolution of the extended system S^{AB}. Both produce the same effect on S^A. Is this type of analogy always possible? We shall show that it is. Furthermore, we will show that for a given evolution $\rho^A \to \rho'^A$, the operator-sum decomposition itself is also not uniquely determined. Many influences on S^A can lead to the same final state. We shall return to the evolution of open systems between their preparation and their measurement in the following chapter. First, we turn to generalised measurements.

Generalised measurements We continue with a particular projective measurement on the second subsystem S^B. For this measurement, we choose the projectors belonging to the computational basis of $\mathcal{H}_2^B$

$$P_+^B = |0^B\rangle\langle 0^B|, \quad P_-^B = |1^B\rangle\langle 1^B| \; . \tag{13.8}$$

The probabilities of occurrence of the results $+$ and $-$ are

$$p_\pm = \mathrm{tr}_B \left[P_\pm^B \rho'^B P_\pm^B \right] \; . \tag{13.9}$$

$\rho'^B = \mathrm{tr}_A \left[\rho'^{AB}\right]$ is here the reduced density operator of the ancilla system S^B after the unitary evolution with U^{AB}. Following a corresponding communication, a selection is carried out at A according to the results $+$ or $-$ obtained at B. We can then interpret $p_\pm$ also with a

view to the subsystem S^A. The $p_\pm$ are at the same time the probabilities that the system S^A is to be found in one of the states

$$\rho'^A \to \tilde{\mu}'^A_\pm = \mathrm{tr}_B \left[P^B_\pm \rho'^{AB} P^B_\pm \right] = \frac{1}{2} \sigma^A_{x,y} \rho^A \sigma^A_{x,y} \tag{13.10}$$

following the measurement. We have made use of Eq. (13.3) in this derivation. The tilde again denotes the fact that the state is not normalised. The normalisation factor would be 2. The overall evolution of S^A is thus $\rho^A \to \rho'^A \to \mu^A_\pm$.

The states $\tilde{\mu}^A_\pm$ of the subsystem S^A after the measurement on S^B were not obtained by projection of the initial state ρ^A. How are the superoperators on $\mathcal{H}^A_2$ which transfer the initial state ρ^A to these two states constructed? To answer this question, we write out the first equation of (13.10) for the measurement result $+$ using Eqs. (13.1) and (13.3):

$$\begin{aligned} \tilde{\mu}'^A_+ &= \mathrm{tr}_B \left[(\mathbb{1}^A \otimes |0^B\rangle\langle 0^B|) U^{AB} (\rho^A \otimes |0^B\rangle\langle 0^B|) U^{AB} (\mathbb{1}^A \otimes |0^B\rangle\langle 0^B|) \right] \\ &= \mathrm{tr}_B \left[\langle 0^B|U^{AB}|0^B\rangle \rho^A \langle 0^B|U^{AB}|0^B\rangle \otimes |0^B\rangle\langle 0^B| \right] \ . \end{aligned} \tag{13.11}$$

We then introduce

$$M^A_+ := \langle 0^B|U^{AB}|0^B\rangle = \frac{1}{\sqrt{2}} \sigma^A_x \ , \quad M^{A\dagger}_+ M^A_+ = \frac{1}{2} \mathbb{1}^A \tag{13.12}$$

and find

$$\tilde{\mu}'^A_+ = M^A_+ \rho^A M^{A\dagger}_+ \ . \tag{13.13}$$

Correspondingly, for the result $-$, we find

$$\tilde{\mu}'^A_- = M^A_- \rho^A M^{A\dagger}_- \tag{13.14}$$

with

$$M^A_- := \langle 1^B|U^{AB}|0^B\rangle = \frac{1}{\sqrt{2}} \sigma^A_y \ , \quad M^{A\dagger}_- M^A_- = \frac{1}{2} \mathbb{1} \, . \tag{13.15}$$

The *measurement operators* $M^A_{+,-}$ on $\mathcal{H}^A_2$ obey the completeness relation

$$M^{A\dagger}_+ M^A_+ + M^{A\dagger}_- M^A_- = \mathbb{1} \ . \tag{13.16}$$

With a similar rearrangement of Eq. (13.9), one can readily confirm that also the measurement probabilities $p_\pm$ can be written with the help of the measurement operators in the form

$$p_+ = \mathrm{tr}_A \left[M^A_+ \rho^A M^{A\dagger}_+ \right] = \mathrm{tr}_A [\tilde{\mu}'^A_+] \tag{13.17}$$

(tr_B and tr_A commute). We find a corresponding result for p_-. The relation $p_+ + p_- = 1$ holds.

We have initially entangled the systems S^A and S^B by means of a unitary evolution. A subsequent projective measurement on the second subsystem S^B leads with the probabilities

p_+ or p_- to the measured results $+$ or $-$. After a selection which depends on the measurement result obtained, the initial system S^A is transformed into one of two well-determined states (13.10). We can interpret this non-projective intervention into S^A as a *generalised selective measurement* on S^A. The measurement results $+$ or $-$ belong to it, with the probabilities p_+ or p_-, respectively, and the non-normalised final states $\tilde{\mu}_+^{'A}$ or $\tilde{\mu}_-^{'A}$. We shall return to this topic in more detail in Sect. 13.3. The results of this measurement on S^A can be formulated completely using the measurement operators $M_\pm^A$, which act only on the state ρ^A of S^A prior to the measurement.

This particular example shows that generalised measurements allow the description of situations which differ clearly from the traditional agenda of quantum measurements, which is oriented around projective measurements. With Eqs. (13.17) and (13.10), we indeed find

$$p_+ = p_- = \frac{1}{2} \,. \tag{13.18}$$

In our example, the probabilities of the two measurement results thus do not depend on the state ρ^A on which the measurement is performed. For physically-relevant generalised measurements, this is however not the case. After having described for the purpose of clarity what is to a certain extent an "extreme case" of a generalised measurement, we shall in the following section discuss a physical example in which the description in terms of a generalised measurement follows in a natural way. The general structure will then be treated in Sect. 13.3.

13.2 The Non-optimal Stern-Gerlach Experiment and Generalised Measurements

13.2.1 The Experimental Setup

The experimental setup In the *Stern-Gerlach experiment* (S-G experiment), spin-$\frac{1}{2}$ objects are produced in a source and after leaving the source in the y direction, they pass through an inhomogeneous magnetic field $\mathbf{B}$. The magnetic field vector lies approximately in the z direction (see Fig. 13.1), and its magnitude is a function of z. The magnetic moments of the objects give rise to an interaction with the magnetic field which in turn causes a polarisation-dependent force to act on the objects. It leads to a deflection of the objects in the state $|\uparrow\rangle$ $(= |0_z\rangle)$ in the positive z direction and of the objects in the state $|\downarrow\rangle$ $(= |1_z\rangle)$ in the negative z direction. For our considerations, this rather idealised description suffices. Details of the calculation of the paths of the objects through the apparatus can be found in textbooks on quantum theory. The deflected objects impact on a detector screen in the $x - z$ plane and produce visible spots there.

The optimal Stern-Gerlach experiment In the optimal S-G experiment, the Schrödinger function $\psi_+(\mathbf{r})$, which describes the behaviour of the polarisation $|\uparrow\rangle$ as a function of position, is localised in a narrow region around the upper path in Fig. 13.1 and the impact points lie only in the upper half-plane ($z > 0$) of the detector screen. Objects with spin $|\downarrow\rangle$ have their state functions $\psi_-(\mathbf{r})$ localised around the lower path in Fig. 13.1) in configuration space and thus

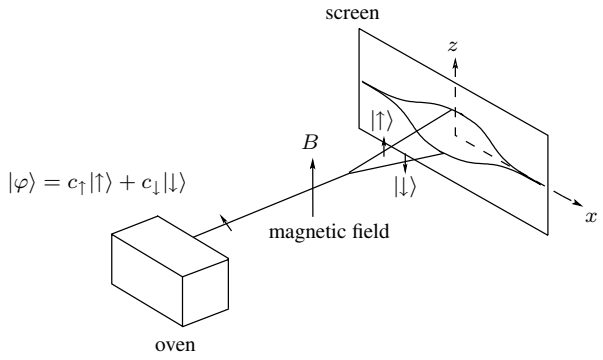


Figure 13.1: The Stern-Gerlach experiment (drawing after [BGL 95]).

impact only in the lower half-plane ($z < 0$) of the screen. The probability distributions for these two cases are shown in Fig. 13.2a. In the optimal S-G experiment, they do not overlap.

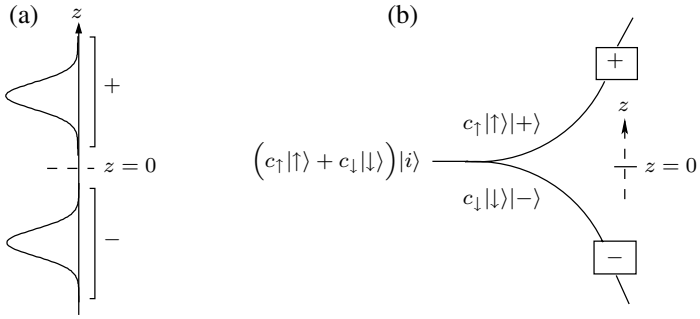


Figure 13.2: The optimal Stern-Gerlach experiment: impact probabilities (a) and the analogous experiment (b). The detectors $+$ and $-$ produce signals for impacts in the upper or in the lower half-plane, respectively.

Simplified description We will consider every impact in the range $z > 0$ uniformly as a measurement result $+1$ and for $z > 0$ as the result -1. The objects are described in terms of states in a product Hilbert space. It is composed of the spin space $\mathcal{H}_2^S$ of the internal degree of freedom, and the space of the outer orbital degrees of freedom. In an ideal S-G experiment, the spin polarisations are measured in the z direction and hence the measured observable is σ_z.

In order to make the computations more clear, we simplify the description and introduce in place of the orbital space a space $\mathcal{H}_2^B$ with the ONB $\{|+\rangle, |-\rangle\}$. The observable $O = |+\rangle\langle+| - |-\rangle\langle-|$ with the measurement outcomes $+1$ and -1 represents an *analogy* to the impacts in the half-planes $z > 0$ and $z < 0$ of the detector screen. We implement it by a signal from a projectively-acting $(+)$ detector and a $(-)$ detector as shown in Fig. 13.2b. The complex deflection process in the S-G experiment has then been reduced to two "orbits" in

our *analogous* experiment. In contrast to the detector screen, our detectors carry out non-destructive measurements. It is thus reasonable to speak of the state of the object after the measurement. We will in the following hence not calculate the S-G experiment; it serves only as the motivation for our assumptions relating to a simpler experiment which exhibits certain analogies to the S-G experiment.

The non-optimal Stern-Gerlach experiment In the non-optimally implemented S-G experiment, the Schrödinger functions $\psi_\pm(\mathbf{r})$ are not strictly localised within the half-planes. The function $\psi_+(\mathbf{r})$ does not vanish in the range $z < 0$ (cf. Figs. 13.3a and 13.3b). The result of this is that for objects with spin $|\uparrow\rangle$, there is a certain probability that the $(-)$ detector will produce a signal, and for those with spin $|\downarrow\rangle$, the $(+)$ detector can respond. The presence of a signal from one detector can thus no longer be taken to be a definite indication of the presence of a particular spin polarisation. In the following, we discuss in detail the various resulting physical situations, which up to now we have described only qualitatively.

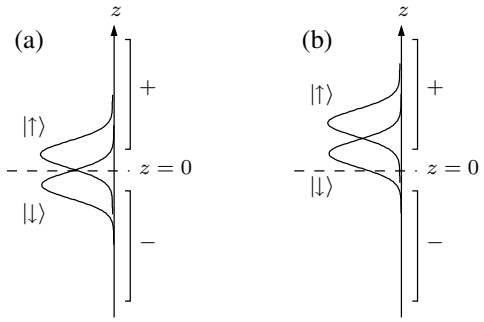


Figure 13.3: The non-optimal Stern-Gerlach experiment. A signal from one of the detectors does not allow us to reach a unique conclusion about the spin polarisation.

13.2.2 An Example of a Generalised Measurement

Before passing through the inhomogeneous magnetic field, the system is in the spin state

$$|\varphi\rangle = c_\uparrow|\uparrow\rangle + c_\downarrow|\downarrow\rangle, \quad |c_\uparrow|^2 + |c_\downarrow|^2 = 1 \tag{13.19}$$

and the orbital state $|i\rangle$. The composite state in $\mathcal{H}_2^S \otimes \mathcal{H}_2^B$ is the incoming product state

$$|\chi\rangle = |\varphi\rangle|i\rangle \ . \tag{13.20}$$

The interaction correlates the states $|\uparrow\rangle$ and $|+\rangle$ as well as the states $|\downarrow\rangle$ and $|-\rangle$ with each other. This leads to entanglement. In the optimal experiment, $|\chi\rangle$ is transformed into the entangled state unitarily:

$$|\chi\rangle \to |\chi'\rangle = c_\uparrow|\uparrow\rangle|+\rangle + c_\downarrow|\downarrow\rangle|-\rangle \ . \tag{13.21}$$

In a projective measurement in the computational basis, the $(+)$ detector responds with the probability p_+ and the spin system is transferred to the state $|\uparrow\rangle$. The corresponding situation holds for the $(-)$ detector.

$$\text{“}+\text{”}: |\varphi\rangle \to |\uparrow\rangle, \quad p_+ = |c_\uparrow|^2 \; ; \qquad \text{“}-\text{”}: |\varphi\rangle \to |\downarrow\rangle, \quad p_- = |c_\downarrow|^2 \; . \tag{13.22}$$

In the optimal analogue experiment, the projective measurement in the orbital space $\mathcal{H}_2^B$ effected by the detectors leads to an indirect projective measurement of the observable σ_z in the spin space. This is shown schematically in Fig. 13.2b.

In a real S-G experiment, in contrast, the two probability distributions no longer lie symmetrically around $z = 0$, but instead are asymmetrically shifted and each protrudes beyond $z = 0$ (see Fig. 13.3). The initial state with the spin state $|\uparrow\rangle$ is transferred in the analogue experiment via the interaction to

$$|\uparrow\rangle|i\rangle \to |\uparrow\rangle(\sqrt{1-p_0}\,|+\rangle + \sqrt{p_0}\,|-\rangle) \tag{13.23}$$

with $0 \le p_0 \le 1$. The corresponding transfer applies to $|\downarrow\rangle$:

$$|\downarrow\rangle|i\rangle \to |\downarrow\rangle(\sqrt{p_1}\,|+\rangle + \sqrt{1-p_1}\,|-\rangle) \tag{13.24}$$

with $0 \le p_1 \le 1$. p_0 is the probability that the $(-)$ detector responds. In the optimal case, the result is $p_0 = 0$ and $p_1 = 0$. The parameters p_0 and p_1 correspond to the shifts of the probability distributions. The normalisation of the state vectors determines the form of the prefactors in Eqs. (13.23) and (13.24).

With equations (13.23) and (13.24), it is at the same time clear how the general incoming state $|\chi\rangle = |\varphi\rangle|i\rangle$ becomes entangled with $|\varphi\rangle$ from Eq. (13.19) due to the interaction:

$$\begin{aligned} |\chi\rangle \to |\chi'\rangle \;\; = \;\; & \{\sqrt{1-p_0}\,c_\uparrow|\uparrow\rangle + \sqrt{p_1}\,c_\downarrow|\downarrow\rangle\}|+\rangle \\ & +\{\sqrt{p_0}\,c_\uparrow|\uparrow\rangle + \sqrt{1-p_1}\,c_\downarrow|\downarrow\rangle\}|-\rangle \; . \end{aligned} \tag{13.25}$$

A measurement again consists of a signal from the $(+)$ detector or the $(-)$ detector. We introduce the measurement operators

$$\begin{aligned} M_+ \;\; := \;\; & \sqrt{1-p_0}\,|\uparrow\rangle\langle\uparrow| + \sqrt{p_1}\,|\downarrow\rangle\langle\downarrow| \\ M_- \;\; := \;\; & \sqrt{p_0}\,|\uparrow\rangle\langle\uparrow| + \sqrt{1-p_1}\,|\downarrow\rangle\langle\downarrow| \; . \end{aligned} \tag{13.26}$$

They obey the completeness relation

$$M_+^\dagger M_+ + M_-^\dagger M_- = \mathbb{1} \; . \tag{13.27}$$

As a generalisation of Eq. (13.22), we then find (compare Fig. 13.4a)

$$\begin{aligned} \text{“}+\text{”}: \quad |\varphi\rangle \to |\varphi_+\rangle \;\; &= \;\; \{\sqrt{1-p_0}\,c_\uparrow|\uparrow\rangle + \sqrt{p_1}\,c_\downarrow|\downarrow\rangle\} \cdot \frac{1}{\text{Norm}} \\ &= \;\; M_+|\varphi\rangle \cdot \frac{1}{\text{Norm}} \, , \qquad (13.28) \\ p_+ \;\; &= \;\; (1-p_0)|c_\uparrow|^2 + p_1|c_\downarrow|^2 = \langle\varphi|M_+|\varphi\rangle \, , \qquad (13.29) \\ \text{“}-\text{”}: \quad |\varphi\rangle \to |\varphi_-\rangle \;\; &= \;\; \{\sqrt{p_0}\,c_\uparrow|\uparrow\rangle + \sqrt{1-p_1}\,c_\downarrow|\downarrow\rangle\} \cdot \frac{1}{\text{Norm}} \, , \\ &= \;\; M_-|\varphi\rangle \cdot \frac{1}{\text{Norm}} \qquad (13.30) \\ p_- \;\; &= \;\; p_0|c_\uparrow|^2 + (1-p_1)|c_\downarrow|^2 = \langle\varphi|M_-|\varphi\rangle \; . \qquad (13.31) \end{aligned}$$

The analogue non-optimal experiment represents a generalised measurement. The analogy with Eqs. (13.12)–(13.17) is apparent.

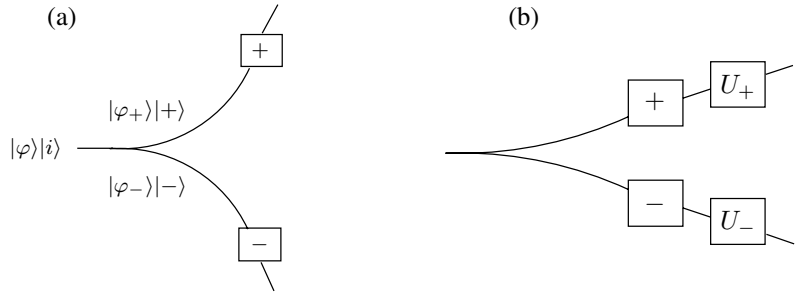


Figure 13.4: (a) A schematic representation of the analogous non-optimal experiment. (b) Extension via unitary influences to a non-minimal measurement.

13.2.3 Unsharp Measurements

In the case of the non-optimal experiment, the projective measurement in the orbital space $\mathcal{H}_2^B$ no longer leads to a projection in spin space. The probabilities p_+ and p_- no longer permit the determination of $|c_\uparrow|^2$ and $|c_\downarrow|^2$. *A generalised measurement is in this sense an unsharp measurement of the spin polarisation*. There are two limiting cases:

(i) $p_0 = p_1 = 0$: the state $|\chi'\rangle$ exhibits the greatest possible degree of entanglement for a given spin state $|\varphi\rangle$. The measurement is a *sharp (or exact)* measurement, since the measurement results lead directly to $|c_\uparrow|^2$ and $|c_\downarrow|^2$. The apparatus is perfectly adjusted. The gain in information about the initial state is in this case maximal. The initial state, on the other hand, is most strongly modified by the measurement.

(ii) $p_0 = p_1 = \frac{1}{2}$: the state $|\chi'\rangle$ is not entangled at all, ($|\chi'\rangle = \frac{1}{\sqrt{2}}|\varphi\rangle \otimes \{|+\rangle + |-\rangle\}$). The measurement is completely *unsharp (or inexact)*. Owing to $p_+ = p_- = \frac{1}{2}$, the measurements permit no conclusions concerning $|c_\uparrow|^2$ or $|c_\downarrow|^2$ to be drawn. The apparatus is completely useless for the determination of these quantities. There is no gain in information. On the other hand, the initial state is not modified by the measurement. We note also that this choice of parameters corresponds to the situation in which in Fig. 13.3 the two curves lie on top of each other and are symmetric with respect to $z = 0$.

When the values of the parameters p_0 and p_1 lie in the neighbourhood of $\frac{1}{2}$, the modification of the spin state is weak. Correspondingly, the gain in information from the measurement is small. The measurement is unsharp. Such measurements are useful if one wishes to follow the time evolution of a state by means of a series of measurements, without perturbing it noticeably (compare Sect. 13.5).

We return once again to the system S^A and the ancilla system S^B as in Sect. 13.1.2. The non-optimal experiment in Sect. 13.2.2 is an example of how the choice of S^B (here the orbital system) and different adjustments of the unitarily-produced entanglement with S^A (here with

the spin system) can produce a continuous spectrum of types of generalised measurements of S^A. For this purpose, a projective measurement is performed on S^B. Is it possible in this way to carry out an arbitrary given non-projective measurement on every system S^A? We shall return to this question in Sects. 13.3.5, 13.4 and in Sect. 16.3.

13.3 Generalised Measurements

13.3.1 What is a Quantum Measurement?

In Chap. 2, the unitary dynamics (transformation dynamics) and the measurement dynamics were introduced. Both types of evolutions are due to *interventions* into the individual quantum systems. In the preceding chapters 13.1 and 13.2, we have seen that there can be non-unitary deterministic evolutions between two measurements and that it is possible to perform non-projective measurements. We wish to describe this more general situation precisely and we will start with the measurements. What kind of interventions are *quantum measurements*? We first describe them in words. Here, we weaken the requirements that are demanded of projective measurements. A mathematical formulation of the measurement postulates will be given in the following section.

An individual quantum measurement is an intervention of a special type into a quantum system, which is carried out by an apparatus (measuring device). This can produce an effect on the measuring device, but need not do so (see null measurement). The effect can be determined by reading out a real number (measurement outcome). At the end of the measurement, the measurement outcome is fixed. The possible measurement outcomes occur as the result of an intervention at random.

In addition, a measurement intervention fulfills the following requirement: If it is specified (for example as a protocol) and the quantum systems result from a preparation procedure which yields a particular state, then the probabilities for the occurrence of the various measurement outcomes are also determined. Since a quantum measurement thus defined is much more general than the projective measurements discussed so far, it is termed a *general measurement*.

How does this greater generality make itself apparent? An immediate repetition of the same measurement (i. e. of the same intervention) on the same quantum system need not lead to the same measurement outcome. In general, therefore, neither does the measurement determine the value of a property (e. g. the energy) which the quantum system had before the measurement, nor does the system obtain a property (e. g. a particular value of the energy) as a result of the measurement. The projective measurement is a special case. We made this clear using the example of the Stern-Gerlach experiment. In contrast to a projective measurement, with a general measurement one thus cannot – always excepting special cases – speak of the measurement of a particular physical quantity such as the energy, the spin polarisation etc.

In a further intervention following the first intervention, a selection according to particular measurement results can be carried out. It is required of a measurement intervention that the quantum systems thus selected be in a well-determined state. They have passed through a preparation procedure which is associated with the particular measurement result. Both interventions taken together are referred to as a *general selective measurement*. Since it is not

determined in advance for an individual quantum system which measurement outcome will be observed, i. e. in which state it will be prepared, one refers with a view to the individual system to a *non-deterministic state evolution*. .

If no selection in terms of measurement outcomes is carried out, the interventions constitute a *general non-selective measurement*. Then, just as in the case of projective measurements without selection, a well-determined, usually mixed state is prepared. It depends only on the state previous to the measurement and on the type of the measurement (measuring device) itself. This case thus represents an intervention with a *deterministic state evolution*.

In quantum mechanics, the term "measurement" can be just as easily misunderstood as the term "state", since they both awaken associations with the meanings of these concepts in classical physics. For quantum systems, there are in general no properties which are predetermined before a measurement and whose values can be obtained from the measurement. Measurement means only to carry out a particular intervention, from which, in contrast to other interventions, information in the form of a value indicated by the measuring device can be read off, but which relates primarily to the outgoing state (the selective measurement prepares the state). We will discuss examples of how one can make use of this information to reach conclusions also about the state before the measurement. For many problems, general measurements are superior to projective measurements. We give examples of this, also. In this chapter, we first consider only the mathematically simplest generalisation of projective measurements, namely generalised measurements, and then discuss general measurements in Sect. 14.3. In the literature, often no distinction is made between general measurements and generalised measurements. We proceed step by step; for this purpose, the distinction is helpful.

13.3.2 Generalised Measurement Postulates

A *generalised selective measurement* is described by a set $\{M_m\}$ of linear *measurement operators*. For every *measurement outcome* m which occurs, one and only one measurement operator M_m is declared. For simplicity, we again assume that the values of the measurement outcomes are discrete. The measurement intervention transforms the state $|\psi\rangle$ of the quantum system into the state

$$|\psi\rangle \to |\psi'_m\rangle = \frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}} \; . \tag{13.32}$$

This transformation and the associated measurement outcome m occur with the probability

$$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle \; . \tag{13.33}$$

Since $M_m^\dagger M_m$ is a positive operator, the condition $p(m) \geq 0$ is fulfilled. In order to conserve the probability interpretation ($\sum_m p(m) = 1$), we must additionally require that the measurement operators obey the completeness relation

$$\sum_m M_m^\dagger M_m = \mathbb{1} \; . \tag{13.34}$$

(Compare the example in Sect. 13.1.2).

The generalisation to density operators as states is obtained with the same justification as in Chap. 4:

$$\rho \to \tilde{\rho}'_m = M_m \rho M_m^\dagger \tag{13.35}$$

$$p(m) = \mathrm{tr}[M_m^\dagger M_m \rho] = \mathrm{tr}[\tilde{\rho}'_m] \,, \qquad \rho'_m = \frac{1}{p(m)} \tilde{\rho}'_m \,. \tag{13.36}$$

In the case of a non-selective measurement, we have

$$\rho \xrightarrow{\text{n.s}} \rho'_{n.s.} = \sum_m M_m \rho M_m^\dagger \,. \tag{13.37}$$

We require of the operators M_m only linearity, aside from the condition (13.34). This guarantees that the relations from Chap. 4 for mixtures and in particular the physically immediately plausible Eq. (4.13) apply here, also. The operators M_m need not be Hermitian. They are in general also not projection operators

$$M_m M_{m'} \neq \delta_{m,m'} M_m \,. \tag{13.38}$$

In particular, the number of measurement operators can be greater than the dimension of the Hilbert space. *The M_m are thus orthogonal projectors only in the special case of projective measurements ($M_m = P_m$).* It can be read off from Eqs. (13.32) and (13.38) that on repeating a measurement, the probability distribution of the measurement outcomes changes. In general, there is no connection to observable operators. The measurement of a property is not associated with a set $\{M_m\}$. In Sect. 14.3, we shall see that the generalised measurements described here (each measurement outcome corresponds to only one measurement operator) are not yet the most general type of measurements.

13.3.3 The Polar Decomposition of a Linear Operator

We first carry out some preliminary mathematical deliberations of which we will then make use in the interpretation of generalised measurements.

The bi-orthogonal expansion of a unitary operator Let U be a unitary operator and $\{|v_i\rangle\}$ an ONB of the Hilbert space. Then the unitary transformation

$$|w_i\rangle = U|v_i\rangle \tag{13.39}$$

again gives rise to an ONB $\{|w_i\rangle\}$. *For a given ONB $\{|v_i\rangle\}$, any unitary operator U can be cast in the form*

$$U = \sum_i |w_i\rangle\langle v_i| \tag{13.40}$$

with the ONB $\{|w_i\rangle\}$ of Eq. 13.39.

The polar decomposition and bi-orthogonal expansion of a linear operator Let L be a linear operator. Then $L^\dagger L$, $LL^\dagger$, $\sqrt{L^\dagger L}$ and $\sqrt{LL^\dagger}$ are positive operators. We begin with the spectral decomposition

$$L^\dagger L = \sum_i \lambda_i |r_i\rangle\langle r_i|, \quad \lambda_i \geq 0 \; . \tag{13.41}$$

$\{|r_i\rangle\}$ is an ONB. The action of L leads to the vectors

$$|m_i\rangle := L|r_i\rangle \; , \tag{13.42}$$

for which

$$\langle m_i|m_i\rangle = \langle r_i|L^\dagger L|r_i\rangle = \lambda_i \tag{13.43}$$

holds. For the index values i for which $\lambda_i \neq 0$, we can normalise the vectors $|m_i\rangle$:

$$|l_i\rangle := \frac{1}{\sqrt{\lambda_i}}|m_i\rangle \; . \tag{13.44}$$

These $|l_i\rangle$ are orthonormal

$$\langle l_i|l_j\rangle = \frac{1}{\sqrt{\lambda_i}\sqrt{\lambda_j}}\langle r_i|L^\dagger L|r_j\rangle = \delta_{i,j} \; . \tag{13.45}$$

We extend the set of these vectors $|l_i\rangle$ to obtain an ONB.

Making use of the ONB $\{|r_i\rangle\}$ and $\{|l_i\rangle\}$, we introduce the unitary operator

$$U := \sum_i |l_i\rangle\langle r_i| \; . \tag{13.46}$$

For U, we find with Eq. (13.41)

$$U\sqrt{L^\dagger L} = \sum_i \sqrt{\lambda_i}|l_i\rangle\langle r_i| \; . \tag{13.47}$$

On the other hand, it follows from Eqs. (13.42) and (13.44) that

$$L|r_i\rangle = \sqrt{\lambda_i}|l_i\rangle \; . \tag{13.48}$$

Since the actions of $U\sqrt{L^\dagger L}$ and L on the basis $\{|r_i\rangle\}$ agree, we have

$$L = U\sqrt{L^\dagger L} \; . \tag{13.49}$$

U is uniquely determined by L, if $\lambda_i \neq 0$ for all i. The relation (13.49) is called the *left-polar decomposition* of the linear operator L. This recalls the analogy with the decomposition $c = e^{i\phi}|c|$ of a complex number c into a magnitude and a phase.

We rewrite Eq. (13.49) again making use of Eqs. (13.41) and (13.46)

$$L = \sum_i \sqrt{\lambda_i}|l_i\rangle\langle r_i|, \qquad \lambda_i \geq 0 \; . \tag{13.50}$$

For every linear operator L, there is a bi-orthogonal decomposition (13.50) *in terms of two ONBs,* $\{|l_i\rangle\}$ *and* $\{|r_i\rangle\}$*, whose construction is given above.* We add without proof the *right-polar decomposition*

$$L = \sqrt{LL^\dagger}U\,. \tag{13.51}$$

Note the different order of the operators under the radical sign. $L^\dagger L$ and $LL^\dagger$ have the same eigenvalues λ_i (compare Eq. (13.41))

$$LL^\dagger = \sum_i \lambda_i |l_i\rangle\langle l_i|\,. \tag{13.52}$$

13.3.4 Minimal Measurements and POVM

The measurement operators M_m of a generalised measurement can always be subjected to a polar decomposition

$$M_m = U_m\sqrt{E_m} \tag{13.53}$$

with

$$E_m := M_m^\dagger M_m\,. \tag{13.54}$$

The E_m, as a result of Eq. (13.34), obey the condition

$$\sum_m E_m = \mathbb{1}\,. \tag{13.55}$$

The positive operators E_m form a *POVM* (positive operator valued measure). They are called *effect operators* or *POVM elements*. We shall return to POVM measurements in Sect. 13.3.5. The non-decomposed measurement operator M_m causes as in (13.35) a transformation into the new state which belongs to the measurement outcome m. The same final state is obtained if one first transforms the initial state with the positive measurement operator $\sqrt{E_m}$ and then carries out the unitary evolution U_m. The probability $p(m)$ from Eq. (13.36) for the occurrence of the measurement outcome m is a function only of E_m:

$$p(m) = \mathrm{tr}[\rho E_m]\,. \tag{13.56}$$

The unitary evolution has no influence on the information obtained from $p(m)$. We can therefore imagine a generalised measurement with the measurement outcome m to be formally decomposed into a measurement dynamics represented by $\sqrt{E_m}$, which is correlated with the probability $p(m)$, and a unitary dynamics U_m, which has no influence on $p(m)$, but does contribute to determining the outgoing state ρ'_m. *For a given POVM, there are arbitrarily many operators* U_m *which lead to the same probability distribution* $p(m)$. Generalised measurements with $U_m = \mathbb{1}$, thus in which the additional unitary influence on the state is absent, are called *minimal measurements*. Projective measurements are minimal.

Minimal measurements on qubits We want to assume that only two measurement results, $+$ and $-$, are possible. Owing to the condition (13.55), we then have

$$E_+ = \mathbb{1} - E_- \tag{13.57}$$

and the POVM operators commute

$$[E_+, E_-] = 0 \; . \tag{13.58}$$

Since the operators are positive, they are also Hermitian and there is an orthonormal basis $\{|0\rangle, |1\rangle\}$ with respect to which both are diagonal:

$$E_+ \;=\; a|0\rangle\langle 0| + p_1|1\rangle\langle 1| \; , \tag{13.59}$$

$$E_- \;=\; p_0|0\rangle\langle 0| + b|1\rangle\langle 1| \; . \tag{13.60}$$

The condition (13.57) is fulfilled by $a = 1 - p_0$ and $b = 1 - p_1$. The positiveness of the operators requires that $0 \leq p_0 \leq 1$ and $0 \leq p_1 \leq 1$. The measurement operators for the associated minimal measurement are then found quite generally in the form

$$M_+ \;=\; \sqrt{E_+} = \sqrt{1-p_0}|0\rangle\langle 0| + \sqrt{p_1}|1\rangle\langle 1| \; , \tag{13.61}$$

$$M_- \;=\; \sqrt{E_-} = \sqrt{p_0}|0\rangle\langle 0| + \sqrt{1-p_1}|1\rangle\langle 1| \; . \tag{13.62}$$

As shown by Eqs. (13.28) and (13.30), the experiment described in Sect. 13.2 is based on these measurement operators. It represents a minimal measurement. A non-minimal measurement would be obtained if there were, as in Fig. 13.4b, an additional unitary dynamics following each detector.

13.3.5 Implementation of a Generalised Measurement by Unitary Transformation and Projection

In Sect. 13.1.2, we saw how one can effect a generalised measurement by extending the system S^A to include an ancilla system S^B, performing an entangling unitary transformation on $\mathcal{H}^A \otimes \mathcal{H}^B$, and carrying out a projective measurement on the system S^B. We wish to show now that every generalised measurement can be implemented in this way.

We extend the system S^A by the ancilla system S^B. The dimension of $\mathcal{H}^A$ is arbitrary. The dimension of $\mathcal{H}^B$ is supposed to be equal to the number of measurement operators. In $\mathcal{H}^B$, we choose an ONB $\{|m^B\rangle\}$ and an arbitrary but fixed state $|0^B\rangle$. We define the linear operator $\hat{U}^{AB}$ on a subspace of $\mathcal{H}^A \otimes \mathcal{H}^B$ by

$$\hat{U}^{AB}|\phi^A, 0^B\rangle = \sum_m M_m^A|\phi^A\rangle \otimes |m^B\rangle \tag{13.63}$$

for arbitrary $|\phi^A\rangle$ from $\mathcal{H}^A$. The measurement operators M_m^A and the basis $\{|m^B\rangle\}$ determine $\hat{U}^{AB}$. For any two vectors $|\phi_1^A, 0^B\rangle$ and $|\phi_2^A, 0^B\rangle$ in this subspace, the operator $\hat{U}^{AB}$

conserves the inner product:

$$\begin{aligned}\langle\phi_1^A,0^B|\hat{U}^{AB\dagger}\hat{U}^{AB}|\phi_2^A,0^B\rangle &= \sum_{m,m'}\langle\phi_1^A|M_{m'}^{A\dagger}M_m^A|\phi_2^A\rangle\langle m'^B|m^B\rangle \\ &= \sum_m\langle\phi_1^A|M_m^{A\dagger}M_m^A|\phi_2^A\rangle \\ &= \langle\phi_1^A|\phi_2^A\rangle = \langle\phi_1^A,0^B|\phi_2^A,0^B\rangle\,. \end{aligned} \tag{13.64}$$

We therefore can apply a mathematical theorem which states that in this case, a unitary extension U^{AB} of $\hat{U}^{AB}$ to the whole space $\mathcal{H}^A\otimes\mathcal{H}^B$ exists (cf. Sect. 13.5). This is important because one can then assume that U^{AB} represents a physically-realisable dynamic evolution of the composite system S^{AB}, which can be described by a suitable Hamiltonian on $\mathcal{H}^A\otimes\mathcal{H}^B$. The effect of the unitary operator U^{AB} on the vectors of the subspace reduces to that of $\hat{U}^{AB}$.

If, following the unitary evolution with U^{AB}, one performs a projective measurement in $\mathcal{H}^B$ using the projection operators

$$P_m^B := \mathbb{1}^A\otimes|m^B\rangle\langle m^B|\,, \tag{13.65}$$

the composite system (after communication and selection) is transferred to the state

$$P_m^BU^{AB}|\phi^A,0^B\rangle\cdot\frac{1}{\text{Norm}} = \frac{M_m^A|\phi^A\rangle|m^B\rangle}{\sqrt{\langle\phi^A|M_m^{A\dagger}M_m^A|\phi^A\rangle}}\,. \tag{13.66}$$

The corresponding probability is

$$p(m) = \langle\phi^A,0^B|U^{AB\dagger}P_m^BU^{AB}|\phi^A,0^B\rangle = \langle\phi^A|M_m^{A\dagger}M_m^A|\phi^A\rangle\,. \tag{13.67}$$

In the last step, we made use of $P_m^B = P_m^{B\dagger}P_m^B$ and of Eq. (13.66). With this, we have shown that *every generalised measurement can be physically implemented by extending the system S^A by an ancilla system S^B. By means of a suitable unitary transformation of the composite system S^{AB}, S^{AB} is transformed into an entangled state. Projection measurements on the ancilla system S^B finally lead to the generalised measurement on S^A.* The operators involved are given in Eqs. (13.63) and (13.65).

13.3.6 Entanglement Distillation by means of Generalised Measurements*

We discuss a further example of the practical significance of generalised measurements. Let us assume the existence of a preparation procedure for the state

$$\rho^{AB} = F|\Psi_-^{AB}\rangle\langle\Psi_-^{AB}| + (1-F)|1^A,1^B\rangle\langle1^A,1^B|\,, \tag{13.68}$$

which we can consider to be a state $|\Psi_-^{AB}\rangle$ which contains the impurity $|1^A,1^B\rangle$. A source emits the maximally-entangled state $|\Psi_-^{AB}\rangle$ with the probability F, where $0<F<1$, and the non-entangled state $|1^A,1^B\rangle$ with the probability $(1-F)$. Using the concurrence introduced

*The sections marked with an asterisk * can be skipped over in a first reading.

in Sect. 11.7, one can show that F is a measure of the entanglement. F at the same time reflects the agreement (fidelity) of ρ with the Bell state $|\Psi_-^{AB}\rangle$

$$F = \langle \Psi_-^{AB} | \rho^{AB}(F) | \Psi_-^{AB} \rangle \,. \tag{13.69}$$

We wish to show that local generalised measurements on the two subsystems following a suitable selection based on classical communication lead to an entanglement distillation. To do so, we use the special case ($p_0 = p_1 =: p$ with $0 < p < 1$) of the general minimal measurement of Eqs. (13.61) and (13.62) with the measurement operators

$$M_- = \sqrt{p}|0\rangle\langle 0| + \sqrt{1-p}|1\rangle\langle 1| \tag{13.70}$$

$$M_+ = \sqrt{1-p}|0\rangle\langle 0| + \sqrt{p}|1\rangle\langle 1| \,. \tag{13.71}$$

If the local measurements on the two subsystems lead to the measurement result "$-$", we keep the system; otherwise, we reject it. A simple computation with the aid of

$$M_-^B M_-^A |\Psi_-^{AB}\rangle = \sqrt{p(1-p)}|\Psi_-^{AB}\rangle \tag{13.72}$$

$$M_-^B M_-^A |1^A, 1^B\rangle = (1-p)|1^A, 1^B\rangle \tag{13.73}$$

shows that after selection, a state of the form (13.68) is present, with fidelity

$$F' = \frac{Fp}{Fp + (1-F)(1-p)} \tag{13.74}$$

instead of F. The probability p_{--} for obtaining the measurement result "$-$" for both subsystems is found to be

$$p_{--} = (1-p^2)\left[Fp + (1-F)(1-p^2)\right] \,. \tag{13.75}$$

We can read off from Eq. (13.74) that the distillation becomes more and more effective because ($F' \nearrow 1$) for $p \nearrow 1$. $p = 1$ is a projective measurement. Why then do we use a generalised measurement? The answer is found in the fact that $p \nearrow 1$ implies that the probability $p_{--} \searrow 0$. In the limiting case of $p = 1$, the measurement would have to be carried out on infinitely many systems in order to find even one completely entangled system after the measurement. In practice, therefore, a compromise is necessary. The entanglement must be sufficiently strong for the purpose at hand, and the distilled states should still occur with a sufficiently high probability.

We have shown that by using generalised measurements on the subsystems with suitable values of p and with classical communication (i. e. by means of LOCC, see Sect. 8.1), the non-local state property of entanglement can be amplified. This is not possible using local projective measurements and classical communication. When the initial entanglement vanishes, no entanglement can be generated by means of LOCC.

13.4 POVM Measurements

13.4.1 Measurement Probabilities and Positive Operators

We discuss measurements from a point of view which is still further reduced in comparison to the above sections. After a measurement intervention, only the measurement outcomes m and

their probabilities of occurrence $p(m)$ are supposed to be known. This is clearly the minimum information which an intervention on a quantum system must deliver in order to be able to speak of a measurement at all. Since the measurement intervention is assumed to have a linear effect on the state, there are linear operators E_m associated with the respective measurement outcomes, with which the probabilities $p(m)$ can be written in the form

$$p(m) = \langle\psi|E_m|\psi\rangle \tag{13.76}$$

$$p(m) = \mathrm{tr}[\rho E_m] \,. \tag{13.77}$$

Due to $p(m) \geq 0$, the E_m must be positive operators which as a result of $\sum_m p(m) = 1$ also must fulfill the condition

$$\sum_m E_m = \mathbb{1} \,. \tag{13.78}$$

The operators E_m are the elements of a decomposition of the unit operator into positive operators. The operators E_m are called *POVM elements* or *effect operators*.

The description of a measurement by a POVM is not restricted to the generalised measurement interventions described in Sect. 13.3.2. *The POVM scheme is valid also in more general measurement situations*, which we will discuss in Sect. 14.3. For many problems, it suffices to reach conclusions about the probability distribution $p(m)$. One then refers to a *POVM measurement*. In the special case of the generalised measurements described in Sect. 13.3.2, we have

$$E_m = M_m^\dagger M_m \,. \tag{13.79}$$

For the physical implementation of a given POVM, it suffices to construct the measurement operator $M_m = \sqrt{E_m}$ as in Sect. 13.3.5. *The projection operators P_m of a projective measurement are the special case of a POVM.* They are called the *projection valued measure*, PVM. In contrast to the PVM, however, the number of POVM elements can be greater than the dimension of the space of states.

13.4.2 A Composite Measurement as an Example of a POVM Measurement

We consider an experimental setup (cf. Fig. 13.5) in which a classical random switch steers the incoming spin-$\frac{1}{2}$ objects in the state $|\psi\rangle$ to different measurement devices for the observables σ_x, σ_y, and σ_z with the probabilities w_x, w_y, and w_z ($w_x + w_y + w_z = 1$). The three detectors each indicate as measurement results the values $+1$ or -1. They carry out projective measurements. The associated projection operators are $P_i(+) = |0_i\rangle\langle 0_i|$ and $P_i(-) = |1_i\rangle\langle 1_i|$, with $i = x, y, z$. All together, there are six possible results, $(i, \pm)$. The corresponding measurement probabilities are

$$p(i,\pm) = w_i \langle\psi|P_i(\pm)|\psi\rangle = \langle\psi|E_i(\pm)|\psi\rangle \tag{13.80}$$

with the six POVM operators

$$E_i(\pm) = w_i P_i(\pm) \tag{13.81}$$

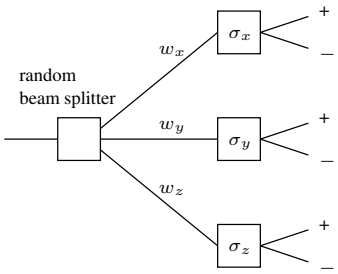


Figure 13.5: A composite measurement.

which obey the completeness relation

$$\sum_i E_i(+) + \sum_i E_i(-) = \sum_i w_i \mathbb{1} = \mathbb{1} \; . \tag{13.82}$$

Due to the random switches which precede the measurements, the projector character is lost. We mention also that the setup represents an *informationally complete measurement*. The statistics reproduced by the measurement probabilities $p(i, \pm)$ permit a unique determination of the initial state even for mixtures (see Problem 13.5). We treat informationally complete measurements in Sect. 13.4.5.

We extend the setup to permit an encompassing measurement by no longer distinguishing between the three detectors. Only the measurement outcome $+1$ or -1 is registered, independently of which detector produced it. Then we find for the remaining two probabilities

$$p(\pm) = \sum_i p(i, \pm) = \langle \psi | E(\pm) | \psi \rangle \tag{13.83}$$

with the operators

$$E(\pm) = \textstyle\sum_i w_i P_i(\pm) \quad , \quad E(+) + E(-) = \mathbb{1} \; . \tag{13.84}$$

We have again performed a POVM measurement.

13.4.3 Can One Distinguish Between Two States with Certainty by a Single POVM Measurement?

Preliminary mathematical considerations An operator E is assumed to be positive. Then it has a spectral representation

$$E = \sum_i \lambda_i |i\rangle\langle i| \; ; \quad \lambda_i \in \mathbb{R} \; , \quad 0 \leq \lambda_i \leq 1 \tag{13.85}$$

with an ONB $|i\rangle$. The action of E upon an arbitrary normalised vector $|\phi\rangle$ is

$$E|\phi\rangle = \sum_i \lambda_i \langle i|\phi\rangle |i\rangle = \sum_i \lambda_i c_i |i\rangle \tag{13.86}$$

with

$$c_i := \langle i|\phi\rangle \ , \quad \sum_i |c_i|^2 = 1 \ . \tag{13.87}$$

With this, we find

$$\langle\phi|E|\phi\rangle = \sum_i \lambda_i |c_i|^2 \ . \tag{13.88}$$

A particular state $|\phi\rangle$ is taken to have the property

$$\langle\phi|E|\phi\rangle = \sum_i \lambda_i |c_i|^2 = 1 \ . \tag{13.89}$$

The non-vanishing c_i have indices i from a set I. For these indices, it follows from Eqs. (13.89) and $0 \leq \lambda_i \leq 1$ that

$$\lambda_{i\in I} = 1 \ . \tag{13.90}$$

Insertion into Eq. (13.86) leads to the result

$$\langle\phi|E|\phi\rangle = 1 \quad \Longleftrightarrow \quad E|\phi\rangle = |\phi\rangle \ . \tag{13.91}$$

The direction $\Leftarrow$ is trivial. In a similar way, one can show that

$$\langle\phi|E|\phi\rangle = 0 \quad \Longleftrightarrow \quad E|\phi\rangle = 0 \ . \tag{13.92}$$

The impossibility of determining a state by a single POVM measurement We assume that we can decide by means of a single measurement whether an individual quantum object is in a state $|\chi\rangle$. Then there must be a measurement with POVM operators E_m in which a particular measurement outcome $\hat{m}$ occurs with certainty when the state $|\chi\rangle$ is present, and with certainty does not occur when *any other state* $|\Theta\rangle \neq |\chi\rangle$ *from the Hilbert space* is present:

$$p_\chi(\hat{m}) = \langle\chi|E_{\hat{m}}|\chi\rangle = 1 \tag{13.93}$$

$$p_\Theta(\hat{m}) = \langle\Theta|E_{\hat{m}}|\Theta\rangle = 0 \ . \tag{13.94}$$

We thus have with Eqs. (13.91) and (13.92):

$$E_{\hat{m}}|\chi\rangle = |\chi\rangle \quad , \quad E_{\hat{m}}|\Theta\rangle = 0 \ . \tag{13.95}$$

With this, we obtain

$$\langle\chi|\Theta\rangle = \langle\chi|E_{\hat{m}}|\Theta\rangle = 0 \ . \tag{13.96}$$

This means that all the other vectors $|\Theta\rangle$ of the Hilbert space must be perpendicular to $|\chi\rangle$. This is however impossible. *It is therefore not possible to determine by means of a single POVM measurement the state of a quantum object before the measurement. Furthermore, this result shows that there is no POVM measurement that would allow us to distinguish between two non-orthogonal states.* This important statement, which is the basis of many heuristic considerations, holds not only for projection measurements, but indeed for arbitrary measurements.

13.4.4 The Advantage of a POVM Measurement for Determining States

For a somewhat modified purpose, in contrast, the application of POVM measurements is more favourable than that of projective measurements. We suppose that we (or a spy) know that a quantum system – is need not be a qubit – has been prepared with equal probabilities in either a state $|1\rangle$ or in a state which is not orthogonal to it, $|2\rangle$. After a single measurement, we want either not to be able to make any statement at all ("don´t know"), or else to know with certainty: "in $|1\rangle$" or "in $|2\rangle$". There must therefore be at least three possible measurement outcomes. The POVM operators are E_1, E_2, and E_3. We construct them in such a way that they have the following property: if the state $|1\rangle$ is present, the measurement outcome 2 will never occur (i. e. $\langle 1|E_2|1\rangle = 0$); but the outcomes 1 and 3 can be found. If the state $|2\rangle$ is present, the outcome 1 never occurs (i. e. $\langle 2|E_1|2\rangle = 0$), but the results 2 and 3 are possible. We can thus draw the following conclusions from a single measurement: if the outcome 1 is obtained, then the state $|1\rangle$ was present. If the result was 2, then $|2\rangle$ was present. No conclusion can be drawn for the outcome 3. This means in terms of POVM operators that

$$E_1 = a_1(\mathbb{1} - |2\rangle\langle 2|) , \tag{13.97}$$

$$E_2 = a_2(\mathbb{1} - |1\rangle\langle 1|) \tag{13.98}$$

and

$$E_3 = \mathbb{1} - E_1 - E_2 . \tag{13.99}$$

It is clearly reasonable to optimise the measurement. To this end, the parameters a_1 and a_2 must be chosen in such a way that the probabilities p which allow a certain conclusion are maximised. Certain conclusions can be drawn only in the cases of the measurement outcomes 1 and 2. The probabilities that these outcomes occur are by construction

$$p = \frac{1}{2}\langle 1|E_1|1\rangle + \frac{1}{2}\langle 2|E_2|2\rangle \tag{13.100}$$

$$= \frac{1}{2}(a_1 + a_2)\left(1 - |\langle 1|2\rangle|^2\right) \tag{13.101}$$

In [Bus 97], it is shown that taking account of the positivity of E_3, through which the parameters a_1 and a_2 are connected, p assumes its maximum value

$$p_{\max} = 1 - |\langle 1|2\rangle| \tag{13.102}$$

when the parameters are chosen to be

$$a_1 = a_2 = \frac{1}{1 - |\langle 1|2\rangle|} . \tag{13.103}$$

Equation (13.103) shows that the probability of being able to make certain statements as the result of a measurement becomes smaller and smaller, the more similar the states $|1\rangle$ and $|2\rangle$ become. Resolving the states by means of a POVM measurement then becomes more and more difficult. For orthogonal states $\langle 1|2\rangle = 0$, $p_{\max} = 1$.

13.4.5 An Informationally-Complete POVM*

A POVM is termed *informationally complete* if for an arbitrary state, a knowledge of the probabilities of all the possible measurement outcomes suffices to determine the state. Thus, only a single measurement device (only one POVM) is used. What conditions must be fulfilled by an informationally-complete POVM for the determination of qubit states?

We saw in Sect. 3.1 that the Pauli operators can be completed to an operator basis on $\mathcal{H}_2$ by $\mathbb{1}$. Every POVM element can be written in the form

$$E_m = a_m \mathbb{1} + b_m \mathbf{n}_m \boldsymbol{\sigma} \; . \tag{13.104}$$

$\mathbf{n}_m$ is here a unit vector in $\mathbb{R}^3$ and a_m and b_m are supposed to be non-negative real numbers, so that E_m and $\mathbb{1} - E_m$ are positive operators. The completeness relation (13.78) leads to the conditions

$$\sum_m a_m = 1 \; , \tag{13.105}$$

$$\sum_m b_m \mathbf{n}_m = 0 \; . \tag{13.106}$$

As we have shown in Sect. 4.4, every density operator ρ on $\mathcal{H}_2$ can be written in the form

$$\rho = \frac{1}{2}(\mathbb{1} + \mathbf{r}\boldsymbol{\sigma}) \tag{13.107}$$

with the Bloch vector $\mathbf{r}$. The probabilities for the measurement outcomes of the POVM measurement are thus

$$p(m) = \mathrm{tr}[\rho E_m] = a_m + b_m \mathbf{n}_m \mathbf{r} \; . \tag{13.108}$$

One can see from this that for the determination of $\mathbf{r}$, the non-vanishing vectors $b_m \mathbf{n}_m$ must span $\mathbb{R}^3$. It then follows together with the linear dependence (13.106) that there must be at least four vectors $\mathbf{n}_m$ $(m = 1, 2, 3, 4)$, and that on the other hand, four vectors which fulfill Eq. (13.106) are sufficient.

We give an example which fulfills Eqs. (13.105) and (13.106): $a_m = b_m = \frac{1}{4}$ and

$$\mathbf{n}_1 = (0, 0, 1), \tag{13.109}$$

$$\mathbf{n}_2 = \left(\frac{2\sqrt{2}}{3}, 0, -\frac{1}{3} \right), \tag{13.110}$$

$$\mathbf{n}_3 = \left(-\frac{\sqrt{2}}{3}, \sqrt{\frac{2}{3}}, -\frac{1}{3} \right), \tag{13.111}$$

$$\mathbf{n}_4 = \left(-\frac{\sqrt{2}}{3}, -\sqrt{\frac{2}{3}}, -\frac{1}{3} \right) \; . \tag{13.112}$$

Insertion into Eq. (13.104) leads to the informationally-complete POVM for qubit states.

*The sections marked with an asterisk * can be skipped over in a first reading.

As we have seen in Sect. 4.4, a measurement of the expectation values of the three different observables σ_x, σ_y, and σ_z allows the determination of the Bloch vector $\mathbf{r}$ and therefore of the state ρ. Here, we have shown that the qubit state ρ can also be identified by measurement of the probabilities $p(m)$ of a single POVM measurement (with only one measurement device) with four possible measurement outcomes m.

13.4.6 Estimating the State Before the Measurement

In classical physics, the state *before* the measurement is determined by carrying out a measurement. In a single quantum measurement, in contrast, when there is no degeneracy, a measured result m allows only the determination of the state *after* the measurement. We will show that on the basis of Bayes' assumption, one can at least make an educated guess as to which state was present before the measurement. The measurement is assumed to be described by the POVM $\{E_m\}$.

We first return to Bayes' theorem as in Eq. (1.102), and make Bayes' assumption (1.103), that all the initial probabilities $p(A_i)$ are the same. In order to make Bayes' theorem clear, we consider the case when the conditional probability $p(B|A_i)$ that B occurs is especially high for a particular event A_j: $p(B|A_j) \gg p(B|A_{i\neq j})$. Let us assume that the result B occurs in fact. Then in this case, A_j was present before with an especially high probability because $p(B|A_{i\neq j})$ is small. We have: $p(A_j|B) \gg p(A_{i\neq j}|B)$. In other words, the event A_j which is followed by B with a particularly high probability is also that which must have been present previously, with the highest probability $p(A_j|B)$, if B in fact occurs. This is the plausible assertion of Bayes' theorem.

We apply this assertion to the quantum-mechanical measurement situation. An individual quantum system is prepared in a pure state. A simple measurement on this system with a measuring device described by a POVM yields the result m. According to Bayes' assumption, the state $|\chi_{pre}\rangle$ which was present before the measurement with the highest probability is characterised by the fact that $p(m)$ is highest for it[2].

The POVM elements are positive operators with the spectral decomposition

$$E_m = \sum_i a_i^{(m)} |r_i^{(m)}\rangle\langle r_i^{(m)}| \,. \tag{13.113}$$

The eigenvectors are presumed not to be degenerate. For $p(m)$, we obtain

$$p(m) = \langle\psi|E_m|\psi\rangle = \sum_i a_i^{(m)} |\langle\psi|r_i^{(m)}\rangle|^2 \tag{13.114}$$

with $|\langle\psi|r_i^{(m)}\rangle| \leq 1$ and $\sum_i |\langle\psi|r_i^{(m)}\rangle|^2 = 1$. We can hence estimate $p(m)$:

$$p(m) \leq a_{\max}^{(m)} \tag{13.115}$$

with $a_{\max}^{(m)} = \max\{a_i^m\}$. The maximum value of $p(m)$ is assumed in Eq. (13.114) for the eigenvector $|\chi_{pre}\rangle = |r_{\max}^{(m)}\rangle$ of E_m belonging to $a_{\max}^{(m)}$. *If no further information is at hand,*

[2] A mathematically exact formulation would have to take into account the fact that the possible states form a continuum (cf. Sect. 13.5).

the best guess for the state before a measurement which yields the result m is the eigenvector $|r^{(m)}_{max}\rangle$ with the largest eigenvalue $a^{(m)}_{max}$ of the POVM element E_m. In a projective measurement ($P_m = |r^{(m)}\rangle\langle r^{(m)}|$), the outgoing state $|r^{(m)}\rangle$ is thus the best guess for the incoming state.

13.5 Complementary Topics and Further Reading

- See Sect. 14.5 for literature on the dynamics of open systems.
- A detailed treatment of the Stern-Gerlach experiment: [BGL 95, Chap. VII].
- The usefulness of unsharp measurements can be seen if one tries to follow the time evolution $|\psi(t)\rangle$ of a state through measurements: [AKS 02], [AKK 04].
- Books on the fundamentals of the theory of quantum operations and of non-projective measurements: [Kra 83], [BGL 95] (see also [HK 69] and [HK 70]).
- The theorem of Neumark describes the implementation of a quantum measurement by a projective measurement on the *composite* system consisting of the original system and its environment: [Per 90].
- Detailed treatments of quantum measurements and POVM: [BGL 95], [Fle 00], [Kon 03].
- The impossibility of determining individual quantum states as well as optimally distinguishing between states: [Bus 97] (and further literature references therein).
- Complementary literature on informationally-complete measurements: [BGL 95], [Aul 00].
- A justification of the results in Sect. 13.4.6 which is not based on Bayes' theorem and Bayes' assumption can be found in [ADK 03].
- A collection of essays on *quantum state estimation*: [PR 04].

13.6 Problems for Chapter 13

Prob. 13.1 [for 13.1]: Confirm Eq. (13.17).

Prob. 13.2 [for 13.2]: For the real Stern-Gerlach experiment in the special case $p_0 = p_1$, determine (using the entropy) the entanglement of the state $|\chi'\rangle$ of Eq. (13.25) and the gain in information due to the measurement.

Prob. 13.3 [for 13.2]: Design a circuit to produce $|\chi'\rangle$ of Eq. (13.25) in the special case that $p_0 = p_1$ making use of the CNOT gate. Which states have to be input?

Prob. 13.4 [for 13.3]: Show that carrying out two measurements in sequence is once again a measurement. Give the corresponding measurement operators.

Prob. 13.5 [for 13.3]: Generalise a result from Sect. 7.7 by proving the following: a non-selective generalised measurement on the subsystem S^A of the bipartite system S^{AB} does not change the state of S^B.

Prob. 13.6 [for 13.3.3]: Prove the right polar decomposition.

Prob. 13.7 [for 13.3.3]: Show, starting from the bi-orthogonal decomposition (13.50) of the linear operator L, that L can always be written in the form

$$L = VDW \ . \tag{13.116}$$

V and W are here unitary operators. D is a positive operator with the eigenvalues λ_i. Hint: rewrite the decomposition by introducing an ONB $\{|a_i\rangle\}$.

Prob. 13.8 [for 13.3.5]: The vectors $\{|n^A, 0^A\rangle\}$ which are generated by an ONB $\{|n^A\rangle\}$ of $\mathcal{H}^A$ are an ONB of a subspace of $\mathcal{H}^A \otimes \mathcal{H}^B$. Let $\hat{U}^{AB}$ be a linear operator defined on the subspace which maps onto $\mathcal{H}^A \otimes \mathcal{H}^B$ and conserves inner product. Then an extension U^{AB} of $\hat{U}^{AB}$ exists, which acts on the whole space as a unitary operator and agrees with $\hat{U}^{AB}$ on the subspace. It could be helpful to employ the dyadic representation and the results of Sect. 13.3.3.

Prob. 13.9 [for 13.3.6]: Determine the concurrence of the state ρ^{AB} in Eq. (13.68).

Prob. 13.10 [for 13.4]: Show that every measurement in which the measurement operators M_m and the POVM elements E_m agree is a projective measurement.

Prob. 13.11 [for 13.4.2]: Show that the composite measurement of Sect. 13.4.2 is informationally complete.

Prob. 13.12 [for 13.4.3]: Prove the assertion (13.92).

Prob. 13.13 [for 13.4.3]: Prove the results (13.102) and (13.103).

Prob. 13.14 [for 13.4.6]: A measuring device described by the measurement operators $\{M_m\}$ carries out a measurement on a quantum object which is in a pure state. The measurement outcome is m. Further information is not available. Estimate the state after the measurement. Take into account that the state before the measurement uniquely determines the state after the measurement. Consider also the special case of a minimal measurement.

© 2007 WILEY-VCH Verlag GmbH & Co.

14 The General Evolution of an Open Quantum System and Special Quantum Channels

Even when the composite system evolves unitarily, the subsystems in general do not follow a unitary evolution. Quantum operations provide a useful approach to the in-out formulation of the dynamic evolution of open systems. We make this clear using the example of quantum channels. The changes of state due to measurements are also in the most general case themselves quantum operations. From this unifying point of view, the scenario and the rules of quantum theory can be formulated anew.

14.1 Quantum Operations and their Operator-Sum Decompositions

14.1.1 Quantum Operations

Dynamic evolution as a quantum operation In connection with quantum channels, teleportation, cryptography, quantum computers and quantum measurements, time evolutions of quantum states occur which are more general than unitary evolutions combined with projective measurements (cf. Chap. 2). In Chapter 13, we encountered examples of general changes of state which can occur when the system is coupled to its environment. In the following, we wish to describe general evolutions of open subsystems without making any assumptions about other systems with which the system under consideration is entangled or with which it interacts dynamically.

In considering an evolution, we always think in terms of a specific external *intervention* and describe its effects on a given density operator. The intervention can for example consist of the passage of quantum objects through a well-defined noisy channel, or carrying out a particular unsharp measurement on them. We thus consider both deterministic as well as non-deterministic evolutions. What is the mathematical structure of the description of the general state evolution of open systems?

The time evolution is described in the Schrödinger representation by a specific mapping of the initial state ρ (with $\mathrm{tr}[\rho] = 1$) onto a final state $\tilde{\rho}'$. Both are formulated as density operators in Liouville space (*in-out scheme*):

$$\rho \to \tilde{\rho}' = \mathcal{E}(\rho) \; . \tag{14.1}$$

The tilde again indicates that also non-normalised states (of trace $\neq 1$) are allowed.

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

For statistical mixtures (see Sect. 4.1) $\rho = p_1\rho_1 + p_2\rho_2$, we require for physical reasons that the action of $\mathcal{E}$ on ρ leads to a statistical mixture of the final states $\mathcal{E}(\rho_i)$ with equal probabilities p_i:

$$\mathcal{E}(\rho) = p_1\mathcal{E}(\rho_1) + p_2\mathcal{E}(\rho_2)\,. \tag{14.2}$$

This establishes the linearity of the mapping. The mapping $\mathcal{E}(\rho)$ must therefore be a superoperator which according to the definition in Sect. 1.2 is linear. A positive superoperator maps positive operators onto positive operators. Since density operators are positive operators, $\mathcal{E}(\rho)$ must be a positive superoperator.

We shall however not require that $\mathcal{E}$ conserve the trace of the density operator. Otherwise, we would have already eliminated the projections which occur in the framework of the projective measurement of an observable. If one writes the generalised measurement from Sect. 13.3 as the action of a superoperator, then it takes the form

$$\rho \to \tilde{\rho}' = \mathcal{E}(\rho) = M_m \rho M_m^\dagger \tag{14.3}$$

and, with $\mathrm{tr}[\rho] = 1$, we have $\mathrm{tr}[\tilde{\rho}] < 1$ (see Sect. 13.3.2 and in particular Eq. (13.33)). We hence require only that the trace of the density operator not increase:

$$\mathrm{tr}[\mathcal{E}(\rho)] \leq 1 \quad \text{when } \mathrm{tr}[\rho] = 1\,. \tag{14.4}$$

Finally, as a third requirement, we demand *complete positivity* of $\mathcal{E}$. The mapping of Eq. (14.1) must not only conserve the positivity of the density operator, but in addition the following must hold: If an arbitrary additional system S^B with the Hilbert space $\mathcal{H}^B$ is added to the system S^A under consideration, and the superoperator $\mathcal{E}^A$ of the evolution of S^A is trivially extended in the form $\mathcal{E}^A \otimes \mathbb{1}^B$ to give an evolution operator of the composite system S^{AB}, then $\mathcal{E}^A \otimes \mathbb{1}^B$ is supposed to again be a positive superoperator on $\mathcal{H}^A \otimes \mathcal{H}^B$. Physically, this means that if only S^A passes through a dynamic evolution, then it must be guaranteed that in the process, the state ρ^{AB} of a composite system S^{AB} again (possibly after normalisation) becomes a density operator ρ'^{AB}. This must be required for reasons of consistency, since it cannot be excluded that the system S^A under consideration is open and hence is a subsystem of a larger system S^{AB}. In this case, via the effect of $\mathcal{E}^A$ on S^A, the composite system S^{AB} could be influenced. Simple examples of this can be found in Eqs. (9.38)-(9.40).

We summarise (see Tab. 14.1): *The general evolution of a quantum system, which in the Schrödinger representation transforms initial states ρ into final states $\tilde{\rho}'$, is a quantum operation $\mathcal{E}$. A quantum operation is a mapping which is described by a superoperator $\mathcal{E}$ which is*

(i) linear,

(ii) does not enlarge the trace,

(iii) and is a completely positive mapping

$$\rho \to \tilde{\rho}' = \mathcal{E}(\rho)\,. \tag{14.5}$$

The final state ρ' is obtained, if necessary, by normalisation

$$\rho' = \frac{\mathcal{E}(\rho)}{\mathrm{tr}[\mathcal{E}(\rho)]}\,. \tag{14.6}$$

We shall assume of the initial state that $\mathrm{tr}[\rho] = 1$, so that condition (ii) means that $\mathrm{tr}[\mathcal{E}(\rho)] \leq 1$.

Not all positive mappings are completely positive. This requirement is a genuine limitation. In Sect. 8.6, we have seen that the transposition T^A is a positive mapping. The partial transposition $T^A \otimes \mathbb{1}^B$ on $\mathcal{H}_2^A \otimes \mathcal{H}_2^B$ or $\mathcal{H}_2^A \otimes \mathcal{H}_3^B$ is however a positive mapping only on separable states.

14.1.2 The Operator-Sum Decomposition of Quantum Operations

Kraus' theorem We formulate here a theorem concerning quantum operations that we will prove in Sect. 16.1 (compare Tab. 14.1). *It is the theorem of the operator-sum decomposition*: *a mapping $\rho \to \tilde{\rho}' = \mathcal{E}(\rho)$ is a quantum operation if and only if it has a decomposition (or a representation)*

$$\mathcal{E}(\rho) = \sum_i K_i \rho K_i^\dagger \tag{14.7}$$

with linear operators K_i which fulfill the condition

$$\sum_i K_i^\dagger K_i \leq \mathbb{1} \tag{14.8}$$

and map the input Hilbert space onto the output Hilbert space. Their dimensions need not be the same. For

$$\sum_i K_i^\dagger K_i = \mathbb{1} \tag{14.9}$$

the operation conserves the trace, otherwise it does not. The decomposition (14.7) is not unique. This theorem is also called *Kraus' theorem (or representation theorem)*. The operators K_i are termed *Kraus operators* or *operation elements*, but also *decomposition operators*. We encountered examples of an operator-sum decomposition in Sects. 13.1.2 and 13.2. Inequalities for operators (such as for example in Eq. (14.8)) were explained in Sect. (1.1.6). $\mathcal{E}(\rho)$ is called a *complete quantum operation* if the equals sign in Eq. (14.8) holds; otherwise $\mathcal{E}(\rho)$ is an *incomplete quantum operation*.

The unitary evolution of the composite system We consider the simple case that a composite system S^{AB} in the non-correlated initial state $\rho^{AB} = \rho^A \otimes |i^B\rangle\langle i^B|$ experiences a unitary evolution with U^{AB}. The result is in general an entangled state ρ'^{AB}. At the same time, a mapping $\rho^A \to \tilde{\rho}'^A$ of the reduced density operator is induced on the subsystem S^A. The Kraus operators for this operation can be read off as follows: ρ^{AB} is transformed into

$$\rho'^{AB} = U^{AB}|i^B\rangle\rho^A\langle i^B|U^{AB\dagger} \,. \tag{14.10}$$

For the state of S^A, this means

$$\rho'^A = \mathrm{tr}_B[\rho'^{AB}] = \sum_n \langle e_n^B|U^{AB}|i^B\rangle\rho^A\langle i^B|U^{AB\dagger}|e_n^B\rangle \,. \tag{14.11}$$

$\{|e_n^B\rangle\}$ is here an ONB of $\mathcal{H}^B$. ρ'^{AB} depends on the initial state $|i^B\rangle$. The CNOT gate is an example. From the unitarity of U^{AB}, for

$$K_n^A := \langle e_n^B|U^{AB}|i^B\rangle \tag{14.12}$$

with $K_n^{A\dagger} = \langle i^B|(U^{AB})^{-1}|e_n^B\rangle$, the completeness relation

$$\sum_n K_n^{A\dagger} K_n^A = \mathbb{1} \tag{14.13}$$

follows, and hence the conservation of the trace. *The deterministic evolution of the subsystem S^A which is induced by U^{AB}*

$$\rho^A \to \rho'^A = \mathcal{E}(\rho^A) = \sum_n K_n \rho K_n^\dagger \tag{14.14}$$

is a trace-conserving quantum operation with the Kraus operators (14.12). It follows immediately that for a given U^{AB}, the operator-sum decomposition is not unique, since $\{e_n^B\}$ can be freely chosen. We will show in Sect. 16.2 that the following converse also holds: *For every trace-conserving quantum operation $\mathcal{E}^A$ on S^A, by extension to a system S^{AB} and choice of an initial state in S^B, a unitary transformation U^{AB} on $\mathcal{H}^{AB}$ can be found, which effects the operation $\mathcal{E}^A$ for the subsystem S^A.* All trace-conserving quantum operations can in this way be reduced to unitary transformations. Unitary transformations are the fundamental trace-conserving transformations.

14.1.3 Simple Quantum Operations

A projection P_m, such as occurs for example in a selective measurement, is a quantum operation $K = P_m$ with $K^\dagger K < \mathbb{1}$. It does not conserve the trace. A unitary transformation U is a trace-conserving operation with $K = U$. If one causes unitary transformations U_i to act upon a system with the probabilities p_i, owing to $\sum_i p_i = 1$, a quantum operation (probabilistic unitary evolution) is again obtained,

$$\mathcal{E}(\rho) = \sum_i p_i U_i \rho U_i^\dagger \ . \tag{14.15}$$

Taking the partial trace is also a quantum operation. To see this, we write the operator K_i^{AB}, which maps from $\mathcal{H}^A \otimes \mathcal{H}^B$ onto $\mathcal{H}^A$, in the following manner (cf. Eq. (7.43)):

$$K_i^{AB}|\psi^{AB}\rangle = K_i^{AB}\left(\sum_j \alpha_j |a_j^A\rangle|e_j^B\rangle\right) = \alpha_i |a_i^A\rangle \ . \tag{14.16}$$

$\{|e_j^B\rangle\}$ is an ONB of $\mathcal{H}^B$ and $|a_j^A\rangle$ is normalised. By its action on the basis vectors of $\mathcal{H}^A \otimes \mathcal{H}^B$, one can confirm that

$$\sum_i K_i^{AB\dagger} K_i^{AB} = \mathbb{1}^{AB} \ . \tag{14.17}$$

Then

$$\mathcal{E}(\rho^{AB}) = \sum_i K_i^{AB} \rho^{AB} K_i^{AB\dagger} \tag{14.18}$$

is a quantum operation. We decompose ρ^{AB} dually in terms of the ONB $\{|l_n^A\rangle|e_i^B\rangle\}$ of $\mathcal{H}^{AB}$, where $\{|l_n^A\rangle\}$ is an ONB of $\mathcal{H}^A$. The action of K_i^{AB} is then given by

$$\mathcal{E}(\rho^{AB}) = \sum_j \langle e_j^B|\rho^{AB}|e_j^B\rangle \; . \tag{14.19}$$

according to Eq. (14.16).

14.1.4 The Ambiguity of the Operator-Sum Decomposition

We allow the following two sets of quantum operations to act on a qubit:

$$\mathcal{E}_i(\rho) \;:=\; \frac{1}{2}\rho + \frac{1}{2}\sigma_i\rho\sigma_i \, , \quad i = x, y, z \tag{14.20}$$

$$\hat{\mathcal{E}}_i(\rho) \;:=\; \langle 0_i|\rho|0_i\rangle |0_i\rangle\langle 0_i| + \langle 1_i|\rho|1_i\rangle \; |1_i\rangle\langle 1_i| \; . \tag{14.21}$$

With $|0_i\rangle\langle 0_i| = (\mathbb{1} + \sigma_i)/2$ and $|1_i\rangle\langle 1_i| = (\mathbb{1} - \sigma_i)/2$, we find after a brief rearrangement agreement for the same index

$$\hat{\mathcal{E}}_i(\rho) = \mathcal{E}_i(\rho) \, . \tag{14.22}$$

We choose a fixed value of i. The action of the two quantum operations is the same, although the physical interpretations of Eqs. (14.20) and (14.21), and the corresponding possible operational implementations are completely different: in Eq. (14.20), either the unitary transformations $\mathbb{1}$ (i. e. no change) or σ_i act upon ρ, each with the probability $\frac{1}{2}$. Equation (14.21) corresponds to a non-selective measurement in the computational basis $\{|0_i\rangle, |1_i\rangle\}$. The Bloch vector $\mathbf{r}$ of ρ' therefore lies parallel to the i axis. This can be readily confirmed with Eq. (14.21) and the relations of Chap. 3. What happens to the Bloch vector when all three Pauli operators σ_x, σ_y, and σ_z, instead of only one Pauli operator σ_i as in $\mathcal{E}_i(\rho)$, act with equal probabilities? We shall return to this question in Sect. 14.4.1.

14.2 The Master Equation

The Lindblad master equation The superoperator $\mathcal{E}$ of a quantum operation connects the initial state with the final state of the dynamic evolution of an open system via an in-out scheme. For practical evaluations, it is useful to be able to follow the time evolution of the density operator by means of a differential equation. Such differential equations are called *master equations*. We will derive a particularly simple type of master equation.

We limit our considerations to particular physical situations in which, in the Schrödinger representation, the state $\rho(t + \mathrm{d}t)$ depends only on the infinitesimally earlier state $\rho(t)$ and on an additive term which is linear in $\mathrm{d}t$,

$$\rho(t + \mathrm{d}t) = \rho(t) + O(\mathrm{d}t) \, . \tag{14.23}$$

The associated quantum operation has a Kraus representation

$$\rho(t+\mathrm{d}t) = \mathcal{E}(\rho(t)) = \sum_i K_i(\mathrm{d}t)\rho(t)K_i^\dagger(\mathrm{d}t)\,. \tag{14.24}$$

The Kraus operators can, like $\mathcal{E}$, themselves depend on the time t. We make use of the ambiguity of the operator-sum decomposition. From the requirement (14.23), it follows with Eq. (14.24) that one of the Kraus operators (e. g. K_1) can be chosen in the form $\mathbb{1} + O(\mathrm{d}t)$. The linear operator $O(\mathrm{d}t)$ can be expanded as in Sect. 1.5 in terms of Hermitian operators

$$K_1 = \mathbb{1} + (R - \frac{i}{\hbar}H)\mathrm{d}t\,, \quad R^\dagger = R\,, \quad H^\dagger = H\,. \tag{14.25}$$

The remaining Kraus operators must be proportional to $\sqrt{\mathrm{d}t}$ for our choice

$$K_i = L_i\sqrt{\mathrm{d}t}\,, \quad i > 1\,. \tag{14.26}$$

The operators L_i are called *Lindblad operators* and are in general neither unitary nor Hermitian. Equation (14.9) leads to the relationship

$$\mathbb{1} = \mathbb{1} + (2R + \sum_{i>1} L_i^\dagger L_i)\mathrm{d}t + O(\mathrm{d}t^2) \tag{14.27}$$

and thus to

$$R = -\frac{1}{2}\sum_{i>1} L_i^\dagger L_i + O(\mathrm{d}t^2)\,. \tag{14.28}$$

Insertion into Eq. (14.24) yields

$$\begin{aligned}\rho(t+\mathrm{d}t) &= \{\mathbb{1} + (R - \frac{i}{\hbar}H)\mathrm{d}t\}\rho(\mathrm{d}t)\{\mathbb{1} + (R + \frac{i}{\hbar}H)\mathrm{d}t\} + \sum_{i>1} L_i\rho(t)L_i^\dagger\mathrm{d}t \\ &= \rho(t) - \{\frac{i}{\hbar}[H,\rho(t)]_- + \sum_{i>1}(L_i\rho(t)L_i^\dagger - \frac{1}{2}[\rho(t), L_i^\dagger L_i]_+)\}\mathrm{d}t \\ &\quad + O(\mathrm{d}t^2)\end{aligned} \tag{14.29}$$

with $[A,B]_- := AB - BA$, $[A,B]_+ := AB + BA$. In the limit $\mathrm{d}t \to 0$, we obtain from this the *Lindblad master equation* in the Schrödinger representation

$$\frac{\mathrm{d}\rho(t)}{\mathrm{d}t} = -\frac{i}{\hbar}[H,\rho(t)]_- + \sum_{i>1}\left(L_i\rho(t)L_i^\dagger - \frac{1}{2}[\rho(t), L_i^\dagger L_i]_+\right)\,. \tag{14.30}$$

The right-hand side of this equation is a superoperator which acts upon ρ. We can write

$$i\hbar\dot{\rho} = \mathcal{L}(\rho) \tag{14.31}$$

in analogy to Eq. (4.10). $\mathcal{L}$ is called the *Lindbladian*.

The Markovian approximation The first term in Eq. (14.30) corresponds to a unitary evolution as we know it from the von Neumann equation (4.9). We can consider ρ to be the reduced density operator of the open subsystem S^A within a closed system S^{AB} ($\rho = \rho^A$). Then the remaining terms describe the influence of the second subsystem S^B on S^A via the Lindblad operators L_i. In a closed system S^A ($L_i = 0$), H would be the Hamiltonian. In the general case, H is not the Hamiltonian of the closed system.

The condition (14.23) is in general only approximately fulfilled. It characterises the *Markovian approximation*, which is applicable in many cases. Whether this approximation is justified for a particular system S^A can be decided only with the help of the microscopic description of S^{AB} and of the interactions between S^A and S^B. We can read off from equation (14.23) that the evolution of the system S^A depends only on the state $\rho(t)$ which is present at the time t and not on earlier states $\rho(t' < t)$. The system has in this sense "forgotten" its history. Although the system S^A is open, its past history does not for example affect it via the second subsystem S^B (i. e. the environment of S^A). In more general situations, events in the past $t' < t$ do influence $\rho^A(t)$ in this manner.

14.3 Completely General Selective Measurements and POVM

Measurement operations instead of measurement operators Measurements are interventions with non-deterministic state evolutions. We have seen in Sect. 13.3 that generalised selective measurements are given by a set $\{M_m\}$ of measurement operators. Each dynamical process belonging to an individual measurement outcome m is a quantum operation in itself (compare Tab. 14.1)

$$\rho \to \tilde{\rho}' = \mathcal{M}_m(\rho) = M_m \rho M_m^\dagger \tag{14.32}$$

with only one Kraus operator M_m. Equation (13.35) shows that

$$p(m) = \mathrm{tr}[\mathcal{M}_m(\rho)] < 1 \tag{14.33}$$

is the probability that the operation $\mathcal{M}_m$ is in fact carried out. In generalised measurements, the operation belonging to a measurement result m thus decreases the trace. For the Kraus operator, we have:

$$M_m^\dagger M_m < \mathbb{1} \ . \tag{14.34}$$

These considerations concerning generalised selective measurements already indicate that there can be a still more general type of measurements in which the operator-sum decomposition of the measurement operations $\mathcal{M}_m$ contains more than one summand. In a generalisation of Eqs. (14.32) and (14.33), we find for completely *general selective measurements* the state $\tilde{\rho}_m^{'}$ after the measurement as the result of a quantum operation with a superoperator $\mathcal{M}_m$:

$$\rho \to \tilde{\rho}'_m = \mathcal{M}_m(\rho) = \sum_i M_{m,i} \ \rho \ M_{m,i}^\dagger \ . \tag{14.35}$$

The range of the index i can depend upon the measurement result m. The probability of occurrence of the measurement results is again

$$p(m) = \text{tr}[\mathcal{M}_m(\rho)] < 1 \; . \tag{14.36}$$

The quantum operation $\mathcal{M}_m$ does not conserve the trace. As a result of $p(m) < 1$, it is not complete

$$\sum_i M_{m,i}^\dagger M_{m,i} < \mathbb{1} \; . \tag{14.37}$$

From $\sum_m p(m) = 1$, we find as a condition on the Kraus operators

$$\sum_{m,i} M_{m,i}^\dagger M_{m,i} = \mathbb{1} \; . \tag{14.38}$$

We show in Section 16.3 that every completely general measurement can be implemented as follows: *The system S^A on which the measurement is to be performed is complemented by an ancilla system S^B. A suitable unitary transformation U^{AB} entangles the two systems. A projective measurement on the ancilla system S^B performs (after communication and selection) a measurement operation $\mathcal{M}_m$ on the system S^A. The probability for the occurrence of $\mathcal{M}_m$ is $p(m)$.*

Non-selective general measurements If no selection according to the measurement result is carried out, the normalised state

$$\rho \xrightarrow{n.s.} \rho'_{n.s.} = \sum_m p(m)\rho'_m \tag{14.39}$$

is obtained. The associated quantum operation is found with Eqs. (14.35) and (14.36):

$$\begin{aligned} \rho'_{n.s.} &= \mathcal{E}(\rho) = \sum_m \text{tr}[\mathcal{M}_m(\rho)] \frac{\mathcal{M}_m(\rho)}{\text{tr}[\mathcal{M}_m(\rho)]} \\ &= \sum_m \mathcal{M}_m(\rho) = \sum_{m,i} M_{m,i}\rho M_{m,i}^\dagger \; . \end{aligned} \tag{14.40}$$

The Kraus operators $M_{m,i}$ fulfill the condition (14.38) for conservation of the trace in the case of a non-selective measurement.

POVM It is found that the POVM measurement introduced in Sect. 13.4 has *a priori* the degree of generality which we have attained for measurements only by introducing the quantum operations in Eqs. (14.35) and (14.36). *Every completely general measurement is associated with a POVM $\{E_m\}$ in terms of the positive operators*

$$E_m := \sum_i M_{m,i}^\dagger M_{m,i} \tag{14.41}$$

which with Eq. (14.38) obey the completeness relation

$$\sum_m E_m = \mathbb{1} \tag{14.42}$$

(cf. Fig. 14.1). The measurement probability is found with (14.36) in the form

$$p(m) = \mathrm{tr}[E_m \rho] \; . \tag{14.43}$$

14.4 Quantum Channels

We can often consider the action of a quantum operation to be the result of passing through a *quantum channel*. We consider two trace-conserving quantum operations. A third example will be discussed in Sect. (15.1)

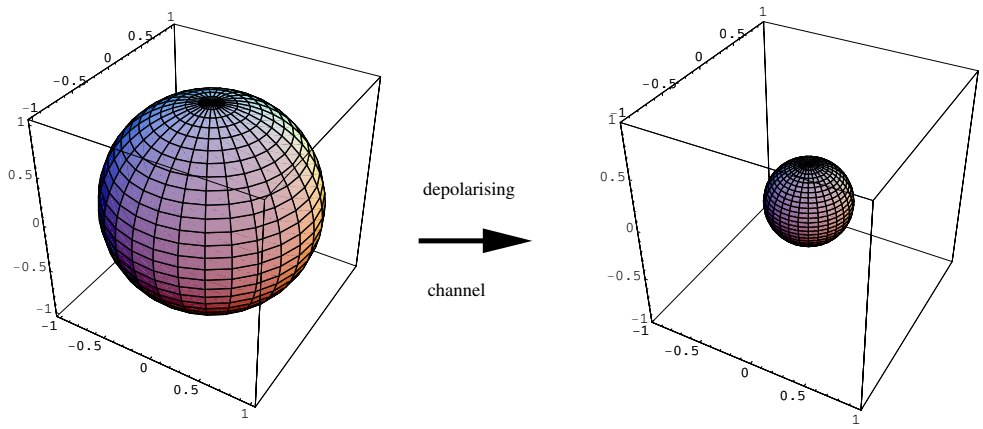


Figure 14.1: The effect of a depolarising channel on the Bloch sphere.

14.4.1 The Depolarising Channel

In order to answer the question posed at the end of Sect. 14.1.4, we introduce as an extension of Eq. (14.20) in addition to $K_0 = \sqrt{1-p}\,\mathbb{1}$ the three Kraus operators

$$K_i = \sqrt{\frac{p}{3}}\,\sigma_i \; , \qquad i = 1,2,3 \; , \tag{14.44}$$

where $0 \leq p \leq 1$. The Kraus operators which characterise the channel are thus determined. $\mathcal{E}(\rho)$ is then

$$\mathcal{E}(\rho) = (1-p)\rho + \frac{p}{3}(\sigma_1\rho\sigma_1 + \sigma_2\rho\sigma_2 + \sigma_3\rho\sigma_3) \, . \tag{14.45}$$

We want to describe the action of $\mathcal{E}$ on the Bloch vector $\mathbf{r}$ of ρ

$$\rho = \frac{1}{2}\left(\mathbb{1} + \mathbf{r}\boldsymbol{\sigma}\right) \ . \tag{14.46}$$

To this end, we choose the coordinate axes in such a way that $\mathbf{r} = r_3\mathbf{e}_3$ and obtain

$$\rho = \frac{1}{2}(\mathbb{1} + r_3\sigma_3) \ . \tag{14.47}$$

With $\sigma_3\sigma_3\sigma_3 = \sigma_3$, $\sigma_1\sigma_3\sigma_1 = -\sigma_3$ and $\sigma_2\sigma_3\sigma_2 = -\sigma_3$, we can again write $\mathcal{E}(\rho)$ in the form of (14.47):

$$\mathcal{E}(\rho) = \frac{1}{2}\left(\mathbb{1} + r_3'\sigma_3\right), \quad r_3' = \left(1 - \frac{4}{3}p\right) r_3, \quad |r_3'| < 1 \ . \tag{14.48}$$

The spin polarisation $\mathbf{r} = \mathrm{tr}[\rho\boldsymbol{\sigma}]$ is reduced. This is termed *depolarisation*. In a centrally-symmetrical manner, all the Bloch vectors are multiplied uniformly by a factor $\left(1 - \frac{4}{3}p\right)$ (see Fig. 14.1). In particular, pure states are thus transformed into mixtures.

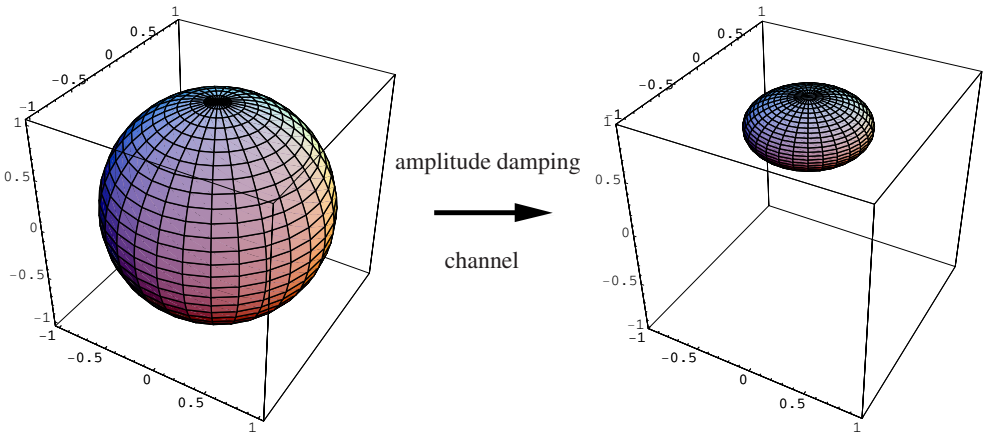


Figure 14.2: The effect of the amplitude damping channel on the Bloch sphere.

14.4.2 Quantum Jumps and Amplitude Damping Channels

We consider a 2-level atom S^A which is passing through an initially empty cavity S^B. If the atom is in an excited state $|1^A\rangle$, it can emit a photon into the cavity with probability p. Then the atom is transferred to its ground state $|0^A\rangle$. The cavity makes the corresponding transition from its initial state $|i^B\rangle = |0^B\rangle$ into the state $|1^B\rangle$.

We compute the Kraus operators for this channel in the matrix representation using Eq. (14.12). Here, we make use of the fact that $|\langle\nu^A, \lambda^B|U^{AB}|\mu^A, \kappa^B\rangle|^2$ gives the probability for the transition $|\mu^A, \kappa^B\rangle \to |\nu^A, \lambda^B\rangle$. In this physical situation, we find

$$
\begin{aligned}
|\langle 0^A, 1^B|U^{AB}|1^A, 0^B\rangle|^2 &= p \\
|\langle 1^A, 0^B|U^{AB}|1^A, 0^B\rangle|^2 &= 1-p \\
|\langle 0^A, 0^B|U^{AB}|0^A, 0^B\rangle|^2 &= 1\,.
\end{aligned}
\tag{14.49}
$$

All the other matrix elements vanish. This leads to the two Kraus operators

$$
\begin{aligned}
K_0^A &= \langle 0^B|U^{AB}|0^B\rangle = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix} \\
K_1^A &= \langle 1^B|U^{AB}|0^B\rangle = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix}\,.
\end{aligned}
\tag{14.50}
$$

They obey

$$
K_0^{A\dagger}K_0^A + K_1^{A\dagger}K_1^A = \mathbb{1}^A\,. \tag{14.51}
$$

The superoperator describes the effect on an arbitrary incoming atomic state ρ^A:

$$
\begin{aligned}
\mathcal{E}(\rho^A) &= K_0^A\rho^A K_0^{A\dagger} + K_1^A\rho^A K_1^{A\dagger} \\
&= \begin{pmatrix} \rho_{00}+p\rho_{11}^A & \sqrt{1-p}\,\rho_{01}^A \\ \sqrt{1-p}\,\rho_{10}^A & (1-p)\rho_{11}^A \end{pmatrix}\,.
\end{aligned}
\tag{14.52}
$$

The probability of finding the atom in its ground state increases. All the other matrix elements of ρ^A decrease. A mixture evolves in the direction of the pure state $|0^A\rangle\langle 0^A|$. For the effects on the surface of the Bloch sphere, see Fig. 14.2. The channel described by Eq. (14.52) is called the *amplitude damping channel*.

14.4.3 An Entanglement-Breaking Channel *

We have already emphasized several times that the entanglement of a state can be reduced when its subsystems propagate through noisy channels to Alice and Bob. We want to demonstrate this explicitly with a simple example. A source generates the maximally-entangled Bell state $|\Psi_-^{AB}\rangle$. The subsystem S^B of the composite system S^{AB} is sent through a depolarising channel (cf. Fig. 14.3). The quantum operation which maps the incoming composite state $\rho^{AB} = |\Psi_-^{AB}\rangle\langle\Psi_-^{AB}|$ onto the outgoing state ρ'^{AB} can be described by the Kraus operators

$$
K_0^{AB} = \sqrt{(1-p)}\mathbb{1}^A\otimes\mathbb{1}^B \qquad K_i^{AB} = \sqrt{\frac{p}{3}}\mathbb{1}^A\otimes\sigma_i^B\,, \quad i=1,2,3 \tag{14.53}
$$

(compare Eq. (14.44)). The operators $\mathbb{1}^A\otimes\sigma_i^B$ transform Bell states into other Bell states (see Eqs. (9.38) - (9.40)). A simple auxiliary calculation yields

$$
\rho'^{AB} = \frac{3-4p}{3}|\Psi_-^{AB}\rangle\langle\Psi_-^{AB}| + \frac{p}{3}\mathbb{1}^{AB}\,. \tag{14.54}
$$

The depolarising channel generates a mixture of the original state and the completely-mixed state $\mathbb{1}^{AB}$. The corresponding situation is found if the outgoing state is another Bell state.

*The sections marked with an asterisk * can be skipped over in a first reading.

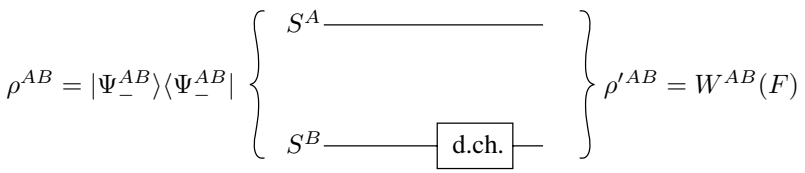


Figure 14.3: The subsystem S^B passes through a depolarising channel (d.ch.). This transforms the composite system S^{AB} into a Werner state $W^{AB}(F)$. When the effects of the channel are sufficiently strong, the entanglement of the initial state is broken.

The Werner state To describe the resulting state, we introduce the notation $F := 1 - p$ and rearrange:

$$\rho'^{AB} = F|\Psi_-^{AB}\rangle\langle\Psi_-^{AB}| + \frac{1-F}{3}(\mathbb{1}^{AB} - |\Psi_-^{AB}\rangle\langle\Psi_-^{AB}|) =: W^{AB}(F)\ . \qquad (14.55)$$

The depolarisation of the subsystem S^B has led to a state which is called a *Werner state*. It is usually denoted by $W^{AB}(F)$. ρ'^{AB} is Bell diagonal. $F = \langle\Psi_-^{AB}|\rho'^{AB}|\Psi_-^{AB}\rangle$ gives the degree of agreement of the Werner state with the initial state. It can be readily shown (see Problem 11.8) that the Werner state is entangled if and only if $F > \frac{1}{2}$ holds. For Kraus operators with $p \geq \frac{1}{2}$, we find $F \leq \frac{1}{2}$. The effect of the channel is so strong (cf. Eq. (14.53)) that the entanglement is broken ("entanglement-breaking channel"). For $F > \frac{1}{2}$, i. e. when there is still a remnant of entanglement in ρ'^{AB}, there are protocols for entanglement distillation which return the Werner states once again arbitrarily close to the initial state $|\Psi_-^{AB}\rangle$ (see Sect. 11.8).

14.5 The Scenario and the Rules of Quantum Theory Revisited

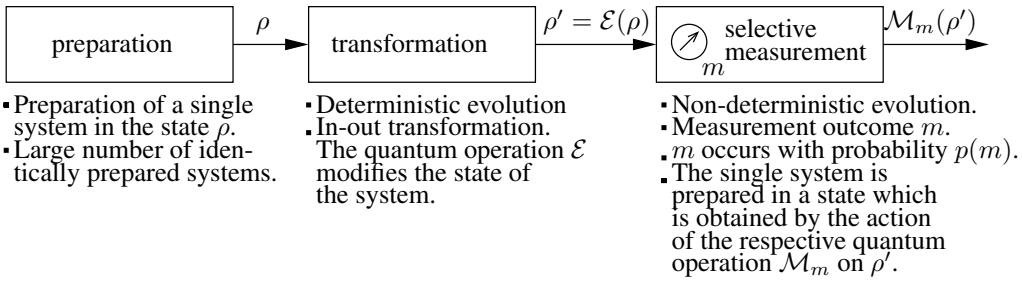

preparation
ρ
transformation
$\rho' = \mathcal{E}(\rho)$
selective measurement
m
$\mathcal{M}_m(\rho')$
▪ Preparation of a single system in the state ρ.
▪ Large number of identically prepared systems.
▪ Deterministic evolution
▪ In-out transformation. The quantum operation $\mathcal{E}$ modifies the state of the system.
▪ Non-deterministic evolution.
▪ Measurement outcome m.
▪ m occurs with probability $p(m)$.
▪ The single system is prepared in a state which is obtained by the action of the respective quantum operation $\mathcal{M}_m$ on ρ'.

Figure 14.4: The scenario of quantum theory in the Schrödinger representation, with the three types of effects upon a quantum system. Without selection with respect to the measured result, a non-selective measurement is performed.

Scenario In the introduction to Sect. 2.1.1, we described the scenario of quantum mechanics. We return here to this topic retrospectively. The quantum system is presumed to be open, i. e. it can be entangled with other quantum systems or it can interact with them. Every experiment in quantum physics, just like every experiment in classical physics, consists of the following sequence: a physical system is affected serially by three types of actions which are caused by three types of apparatus (compare Fig. 14.4):

1. A *preparation apparatus* prepares the system, which can be a subsystem, in a certain state.

2. A *transformation apparatus* acts on the system and changes its state in a *deterministic* fashion (e. g. through the influence of an external potential, interaction with another system, mutual interactions of subsystems). Several transformations can be "passed through" in sequence. In particular, there may not be any transformation apparatus present at all. The effect on the system has the result that when measurements are performed on it (compare point 3), the probabilities for the occurrence of particular measurement results are changed in a deterministic manner. In this sense, a new preparation of the system occurs. This can be formulated in the Schrödinger representation or in other representations.

3. Finally, in the *measuring device* a *non-deterministic* effect occurs, which makes it possible for the measurement outcome to be read out of the measuring device in the form of a real number. For a given state, the probabilities of the measurement outcomes are fixed (probabilistic, irreversible evolution). It is possible that the system is not destroyed by the measurement and that a selection can be carried out in terms of the measurement outcomes. Then the measurement device acts upon the incoming state of the system like an apparatus which prepares different states, depending on the measurement outcome.

Rules To simplify the description, we have interpretively assumed that individual physical systems exist. In many cases, the attempt to explain the experimental results in terms of classical physics fails, while the use of (non-relativistic) quantum theory is successful. Then, in the Schrödinger representation, the following rules are applied, which one can also consider to be an extended version of the postulates in Sect. 2.1.4 (see Tab. 14.1):

- To each preparation procedure, a state is associated. The state is that mathematical object which permits the probability of occurrence of the various measurement outcomes to be predicted for all types of measurements which can be carried out on the correspondingly prepared quantum system. Probability is here usually interpreted as the limiting value of the relative frequency of occurrence. States are density operators ρ on a Hilbert space.

- Transformation is a quantum operation. It is described in the Schrödinger representation by a linear, completely positive superoperator $\mathcal{E}$ on the Liouville space of the operators

$$\rho \to \tilde{\rho}' = \mathcal{E}(\rho) \tag{14.56}$$

(quantum operation), which conserves the trace. With $\mathcal{E}$, an equivalent master equation for $\rho(t)$ can be formulated.

Table 14.1: Interventions which lead to a deterministic or a non-deterministic (probabilistic) state evolution. The trace of operators with a tilde is not equal to one.

	Interventions with a deterministic evolution of state $\Longleftrightarrow$ no measurement or non-selective measurement	**Interventions with a non-deterministic evolution of state $\Longleftrightarrow$ general selective measurement**	
		With information about the final state	Without information about the final state
Quantum operations	The quantum operation $\mathcal{E}$ is trace-conserving and will be carried out with certainty. $\rho \to \rho' = \mathcal{E}(\rho)$ $\mathrm{tr}[\rho'] = \mathrm{tr}[\rho] = 1$	The quantum operation $\mathcal{M}_m$ is not trace-conserving and will be carried out with probability $p(m) < 1$. The respective measurement outcome m can be read off. Selection leads to the state $\rho \to \tilde{\rho}'_m = \mathcal{M}_m(\rho)$. $p(m) = \mathrm{tr}[\tilde{\rho}'_m] = \mathrm{tr}[\mathcal{M}_m(\rho)] < 1,\ \sum_m p(m) = 1$	A positive operator E_m is attributed to each measurement outcome m. The set $\{E_m\}$ is a POVM. $p(m) = \mathrm{tr}[E_m \rho]$ $\sum_m E_m = \mathbb{1}$
Kraus representation	$\mathcal{E}(\rho) = \sum_i K_i \rho K_i^\dagger$ $\sum_i K_i^\dagger K_i = \mathbb{1}$ complete quantum operation $\mathcal{E}$	$\mathcal{M}_m(\rho) = \sum_i M_{m,i} \rho M_{m,i}^\dagger$ $\sum_i M_{m,i}^\dagger M_{m,i} < \mathbb{1},\ \sum_{m,i} M_{m,i}^\dagger M_{m,i} = \mathbb{1}$ incomplete quantum operation $\mathcal{M}_m$	If the quantum operations $M_{m,i}$ are known: $E_m = \sum_i M_{m,i}^\dagger M_{m,i}$
Kraus representation for special cases	Non-selective measurement $\rho \to \rho'_{n.s.} = \mathcal{E}(\rho) = \sum_{m,i} M_{m,i} \rho M_{m,i}^\dagger$ $\mathrm{tr}[\rho'_{n.s.}] = 1$ $\sum_{m,i} M_{m,i}^\dagger M_{m,i} = \mathbb{1}$	Generalised measurement ($i = 1$) $\rho \to \tilde{\rho}'_m = M_m \rho M_m$ $p(m) = \mathrm{tr}[\tilde{\rho}'_m] = \mathrm{tr}[M_m^\dagger M_m \rho] < \mathbb{1}$ $M_m^\dagger M_m < \mathbb{1},\ \sum_m M_m^\dagger M_m = \mathbb{1}$	Generalised measurement $E_m = M_m^\dagger M_m$
	Unitary evolution (i=1, K=U) $U^\dagger U = \mathbb{1}$ $\rho \to \rho' = \mathcal{E}(\rho) = U \rho U^\dagger$ $\mathrm{tr}[\rho'] = 1$	Projective measurement ($i = 1, M_m = P_m$) $P_m^\dagger = P_m,\ P_m P_m = P_m,\ \sum_m P_m = \mathbb{1}$ $\rho \to \tilde{\rho}'_m = P_m \rho P_m$ $p(m) = \mathrm{tr}[\tilde{\rho}'_m] = \mathrm{tr}[P_m \rho] < 1$	PVM $E_m = P_m$
A possible implementation	The system S^A, which is in the state ρ^A, is extended by an ancilla system S^B to form the composite system S^{AB}. An appropriate unitary transformation U^{AB} of the composite system S^{AB} prepares the initial system S^A in the state $\mathcal{E}(\rho^A)$.	The system S^A, which is in the state ρ^A, is extended by an ancilla system S^B to form the composite system S^{AB} The composite system S^{AB} is unitarily transformed in an appropriate way. A subsequent projective measurement of the ancilla system prepares the subsystem with the probability $p(m)$ in the state $\mathcal{M}_m(\rho^A)$.	

- In a selective measurement, the quantum state ρ, depending on the measurement outcome m, is prepared in the state

$$\rho \to \tilde{\rho}'_m = \mathcal{M}_m(\rho) \; . \tag{14.57}$$

Normalisation leads to the new density operator ρ'_m, where $\mathrm{tr}[\rho'_m] = 1$. The measurement interventions corresponding to the different measurement outcomes are represented by quantum operations $\mathcal{M}_m$ which reduce the trace. The set $\{\mathcal{M}_m\}$ is specific to the particular measuring device. The probability of occurrence of the measurement outcome m is (with the convention $\mathrm{tr}[\rho] = 1$)

$$p(m) = \mathrm{tr}[\tilde{\rho}'_m] < 1 \, , \quad \sum_m p(m) = 1 \; . \tag{14.58}$$

- A non-selective measurement is a transformation.

- We note an implication: the measurement operations can be subjected to an operator-sum decomposition

$$\mathcal{M}_m = \sum_i M_{m,i} \rho M_{m,i}^\dagger \, . \tag{14.59}$$

We use the Kraus operator $M_{m,i}$ to construct the effect operators

$$E_m := \sum_i M_{m,i}^\dagger M_{m,i} \, . \tag{14.60}$$

With them, the measurement probabilities $p(m)$ can be written in the form

$$p(m) = \mathrm{tr}[E_m \rho] \; . \tag{14.61}$$

As an expression of $\sum_m p(m) = 1$, we have

$$\sum_m E_m = \mathbb{1} \, . \tag{14.62}$$

The positive operators E_m form a POVM.

- From an abbreviated approach, which aims only at knowledge of the probability distribution $p(m)$, one can characterise a measurement directly without reference to the Kraus operators by giving the POVM $\{E_m\}$ (POVM measurement). Equations (14.61) and (14.62) are valid. Different measurements can lead to the same POVM.

- We have assumed that classically-describable systems and two different dynamics are fundamental to the scenario. This represents no limitation of generality. We do not assert that a different description from that with two distinct types of physical systems (classical and quantum-mechanical) and two distinct dynamics is impossible. In the next chapter, we shall discuss the question as to what extent the introduction of classical concepts on a fundamental level becomes superfluous owing to entanglement and decoherence, so that

the measurement dynamics can be reduced completely to the transformation dynamics. In the case that this research programme is successful and yields all of our results, the rules for the measurement dynamics used above would then acquire the character of a phenomenological description. This represents a pragmatic approach in which one deals for brevity with classical concepts and with two different dynamics as if they were fundamental.

14.6 Complementary Topics and Further Reading

- Quantum operations and operator-sum decompositions: [HK 69], [HK 70], [Lud 83], [Kra 83], [BGL 95], [Sch 96b].
- Description of the dynamics of open systems using master equations, and the Markovian approximation: [Dav 76], [MW 98], [Car 99], [BP 02].
- An axiomatic approach to quantum theory which starts from the scenario described in Sect. 14.5 can be found in [Har 01a], [Har 01b].
- Refer also to the literature references for Chap. 13.

14.7 Problems for Chapter 14

Prob. 14.1 [for 14.1.2]: Show that every quantum operation $\mathcal{E}$ which acts on the density operator of a qubit can be written in the form

$$\mathcal{E}(\rho) = \sum_{i,j=0}^{3} a_{ij}\sigma_i\rho\sigma_j\,, \quad \sigma_0 := \mathbb{1}\,, \tag{14.63}$$

where $a_{ij} = a_{ij}^*$.

Prob. 14.2 [for 14.1.4]: Show by an explicit computation that the Bloch vector of ρ' from Eq. (14.22) lies parallel to the i axis.

Prob. 14.3 [for 14.4]: Calculate the effect of an amplitude damping channel on the density operator

$$\rho = \frac{1}{2}(\mathbb{1} + \mathbf{r}\boldsymbol{\sigma})\,. \tag{14.64}$$

How is the Bloch vector $\mathbf{r}$ modified?

© 2007 WILEY-VCH Verlag GmbH & Co.

15 Decoherence and Approaches to the Description of the Quantum Measurement Process

Decoherence is a parasitic effect for applications, but it is an important approach for the understanding of the quantum measurement process. The decoherence caused by scattering gives an indication that environmentally-induced decoherence could play a central role in the measurement process. The formation of the classical world and the classical behaviour of Schrödinger's cat are further examples of the importance of decoherence. The question as to whether the quantum measurement problem has already been solved today will be treated at the end of this chapter.

15.1 Channels which Produce Decoherence

15.1.1 The Phase Damping Channel

Scattering as a simple example We consider a qubit system S^A on which a quantum system S^B is scattered. We will greatly simplify the scattering process for our description (see Fig. 15.1). To this end, we assume that there exist two orthonormal states $|0^A\rangle$ and $|1^A\rangle$ of the system S^A, which are not modified by the scattering (stable states). The system S^B arrives in the state $|i^B\rangle$. On scattering by the state $|0^A\rangle$ $\big($or $|1^A\rangle\big)$, the scattered system S^B is transformed asymptotically into the state $|0^B\rangle$ $\big($or $|1^B\rangle\big)$. The three states of S^B are supposed to form an ONB in $\mathcal{H}_3^B$. In addition, we want to allow the possibility that with the probability $1-p$, the system S^B is not scattered at all and therefore remains in the state $|i^B\rangle$. One can consider $|0^A\rangle$ and $|1^A\rangle$ to be for example two energy levels, and $|i^B\rangle$, $|0^B\rangle$, and $|1^B\rangle$ to be three momentum states ("orbits").

The scattering is a unitary process of the composite system S^{AB}. The operator which carries the states in the incoming region over to those in the outgoing region has the properties

$$\hat{U}^{AB}|0^A, i^B\rangle = \sqrt{1-p}\,|0^A, i^B\rangle + \sqrt{p}\,|0^A, 0^B\rangle \tag{15.1}$$

$$\hat{U}^{AB}|1^A, i^B\rangle = \sqrt{1-p}\,|1^A, i^B\rangle + \sqrt{p}\,|1^A, 1^B\rangle\ . \tag{15.2}$$

We thus know its effect on parts of the ONB of $\mathcal{H}_2^A \otimes \mathcal{H}_3^B$ and we can extend this operator to obtain a unitary operator on $\mathcal{H}_2^A \otimes \mathcal{H}_3^B$.

$$\begin{aligned} U^{AB} &= \Big(\sqrt{1-p}\,|0^A, i^B\rangle + \sqrt{p}\,|0^A, 0^B\rangle\Big)\,\langle 0^A, i^B| \\ &\quad + \Big(\sqrt{1-p}\,|1^A, i^B\rangle + \sqrt{p}\,|1^A, 1^B\rangle\Big)\,\langle 1^A, i^B| + \text{rest}\ . \end{aligned} \tag{15.3}$$

Further dual combinations with $\langle 0^A, i^B|$ or $\langle 1^A, i^B|$ do not occur in the rest.

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

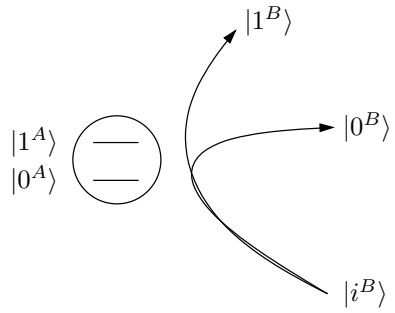


Figure 15.1: Scattering on a 2-level system.

The Kraus operators which belong to the quantum operation that acts on the subsystem S^A can be read off directly according to Eq. (14.12):

$$K_i^A = \langle i^B|U^{AB}|i^B\rangle = \sqrt{1-p}\,\mathbb{1}^A = \sqrt{1-p}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{15.4}$$

$$K_0^A = \langle 0^B|U^{AB}|i^B\rangle = \sqrt{p}\,|0^A\rangle\langle 0^A| = \sqrt{p}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \tag{15.5}$$

$$K_1^A = \langle 1^B|U^{AB}|i^B\rangle = \sqrt{p}\,|1^A\rangle\langle 1^A| = \sqrt{p}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\,. \tag{15.6}$$

The condition

$$K_i^{A\dagger}K_i^A + K_0^{A\dagger}K_0^A + K_1^{A\dagger}K_1^A = \mathbb{1} \tag{15.7}$$

for Kraus operators is fulfilled. The quantum operation which describes the modification of the system S^A if it were in the state ρ before the scattering is given by

$$\rho^A \to \rho'^A = \mathcal{E}(\rho^A) = K_i^A\rho^A K_i^A + K_0^A\rho^A K_0^A + K_1^A\rho^A K_1^A\,. \tag{15.8}$$

In the computational basis, it takes the form

$$\rho'^A = \mathcal{E}(\rho^A) = \begin{pmatrix} \rho_{00}^A & (1-p)\rho_{01}^A \\ (1-p)\rho_{10}^A & \rho_{11}^A \end{pmatrix}\,. \tag{15.9}$$

If scattering occurs with certainty ($p = 1$), the non-diagonal elements of the density matrix of S^A vanish in the computational basis. For $0 \le p < 1$, the non-diagonal elements decrease by the factor $(1-p)$ with each additional individual quantum system S^B which is scattered on S^A. Scattering causes decoherence.

The phase damping channel In the limiting case of $p = 1$ (perfect scattering), the scattering process produces a marking with orthogonal marker states $|0^B\rangle$and $|1^B\rangle$. A pure state is transformed as in Eq. (15.3) into an entangled state

$$\left(c_0|0^A\rangle + c_1|1^A\rangle\right)|i^B\rangle \to c_0|0^A, 0^B\rangle + c_1|1^A, 1^B\rangle\,. \tag{15.10}$$

The pure state of the subsystem S^A becomes a mixture. The phase relation between the summands of the pure state, and thus the coherence, i. e. the capability of interference, are lost for S^A. This is decoherence via marking, which we have already described in Sect. 8.7. The system S^A passes through a quantum channel for $0 \leq p \leq 1$ during the scattering, which is called the *phase damping channel*. The effect of this channel on arbitrary mixtures can be most simply demonstrated on the Bloch sphere. One can show that the Bloch vectors along the z axis remain unchanged. All the other points move for $p \neq 1$ in many scattering cases towards the z axis, whereby the rotational symmetry around the z axis remains (compare Fig. 15.2).

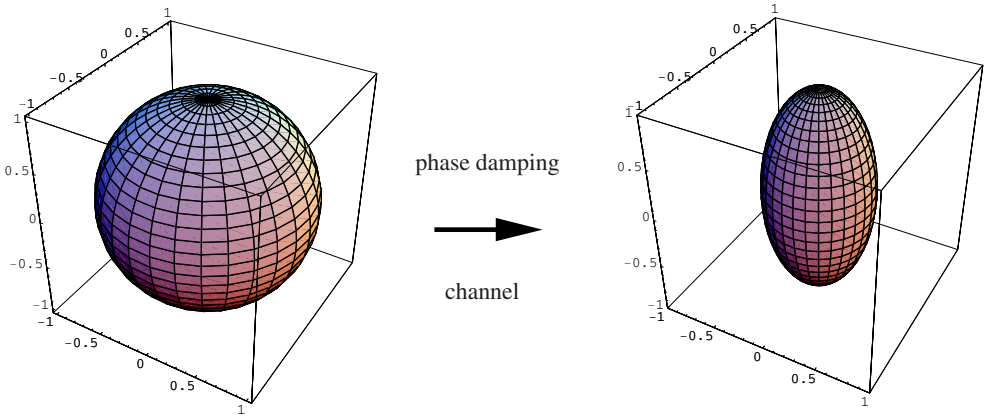


Figure 15.2: Effects of the phase damping channel on the Bloch sphere.

15.1.2 Scattering and Decoherence

Characteristic properties We want to emphasize some properties of the scattering process described above which we shall encounter again to some extent in more general situations in which decoherence plays a role.

a) First of all, we must keep in mind that the fundamental process which the composite system passes through is a unitary evolution U^{AB} that is based upon the particular dynamics of the interaction. A pure state of S^{AB} is transformed into a pure state of S^{AB}. Information is not lost in this process.

b) As a result of this unitarity, the scattering process is reversible. If S^A was initially in a pure state, then by means of a suitable "reflection" (a reversal of the motion and dynamics) of the scattering products, the pure state can again be produced and the decoherence of S^A can be completely reversed. As we have already seen in Sect. 9.1, we have transferred information by the transition (15.10) into the correlations, whence it can in principle again be recalled. This requires a non-local process, which in the case of scattering can not be realised in practice. This last point is important.

c) It is essential that through the interaction, a basis $(|0^A\rangle, |1^A\rangle)$ of stable states is defined in $\mathcal{H}_2^A$, which cannot be changed by the dynamic influence (cf. Eq. (15.9))

$$\mathcal{E}(|0^A\rangle\langle 0^A|) = |0^A\rangle\langle 0^A|\,, \quad \mathcal{E}(|1^A\rangle\langle 1^A|) = |1^A\rangle\langle 1^A|\,. \tag{15.11}$$

d) These eigenstates of σ_z mark the position of the z axis onto which the Bloch sphere shrinks (see Fig. 15.2). Compared with the characterisation of the decoherence via the vanishing of the non-diagonal elements of the density operator, this is a basis-independent characterisation.

e) Due to c), one can already speculate that we have described elastic scattering. If the states $|0^A\rangle$ and $|1\rangle$ belong to energy levels E_0^A and E_1, then with the Hamiltonian

$$H^A = E_0^A|0^A\rangle\langle 0^A| + E_1|1^A\rangle\langle 1^A| \tag{15.12}$$

it can readily be shown that the expectation value $\bar{E}$ of the energy remains unchanged

$$\bar{E}^A = \mathrm{tr}[\rho^A H^A] = \mathrm{tr}[\rho'^A H^A] = \bar{E}'^A\,. \tag{15.13}$$

There is no dissipation.

f) If the scattering is not perfect ($p \neq 1$), the multiple repetitions with new incoming particles will lead to a stepwise increasing decoherence, which drives the Bloch vectors towards a position orthogonal to the z axis.

g) In the limiting case, the density operator

$$\rho_{lim}^A = |c_0|^2|0^A\rangle\langle 0^A| + |c_1|^2|1^A\rangle\langle 1^A| \tag{15.14}$$

results. It is formally the same as the density operator of the statistical mixture in which the system is found in the level state $|0\rangle$ or in $|1\rangle$ with the probability $|c_0|^2$ or $|c_1|^2$, respectively. However, an ignorance interpretation is not possible.

h) The overall process is a preparation procedure for ρ_{lim}^A.

15.1.3 The Phase Flip Channel

We have already seen that there are many operator-sum decompositions for a given quantum operation. Physically, this means that for the same initial state, different dynamic processes can lead to the same final state. We can obtain the quantum operation $\mathcal{E}(\rho)$ as in Sect. 15.1.1, and thus decoherence, also through a *phase flip channel*, which likewise destroys ordered phase relations. In this channel, a dynamic acts with the probability w on the initial state ρ and causes a unitary transformation σ_z – and thus a phase flip. The quantum operation hence has the two Kraus operators

$$K_+ = \sqrt{w}\,\sigma_z \quad , K_- = \sqrt{1-w}\,\mathbb{1} \tag{15.15}$$

with

$$K_+^\dagger K_+ + K_-^\dagger K_- = \mathbb{1}\,. \tag{15.16}$$

The operation is given by

$$\begin{aligned} \rho \to \rho' &= w \begin{pmatrix} \rho_{00} & -\rho_{01} \\ -\rho_{10} & +\rho_{11} \end{pmatrix} + (1-w) \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} \\ &= \begin{pmatrix} \rho_{00} & (1-2w)\rho_{01} \\ (1-2w)\rho_{10} & \rho_{11} \end{pmatrix} \end{aligned} \tag{15.17}$$

in the eigenbasis of σ_z. For $w = \frac{1}{2}$, total decoherence results. *A channel with a random phase flip has an effect which is equivalent in its action to that of a phase damping channel.*

15.2 Environment-Induced Decoherence

15.2.1 The Formation of the Classical World

The programme In Sect. 2.1.1, we described the action of a two-slit experiment on the one hand for balls (i. e. for classical objects) and on the other hand for atoms (i. e. for quantum objects) and emphasized the differences in the experimental results. The theoretical explanation was given in the one case within the framework of classical physics and in the other within the quantum theory in the version which was summarised again in Sect. 14.5. The two-slit experiments raise the following question: what is observed on going from electrons and atoms through viruses and molecules towards larger and larger objects (see Sect. 15.6), finally ending up with tennis balls? Can quantum mechanics, as we have described it, be applied unchanged to explain the observations in the different size ranges? Then classical physics would finally be derivable from quantum mechanics without any additional theories. Or is there a completely new physics which appears at some intermediate size range and requires a new theory?

The question of how classical physics arises from the quantum theory (the quantum-to-classical transition) will certainly not be answered in a completely satisfactory way in the near future. It is, however, reasonable to discuss approaches to the solution of parts of the problem. In this way, we can gain a feeling for what quantum mechanics in the form that we have described can account for and what not. While in Chap. 10 we attempted unsuccessfully to understand quantum physics on the basis of classical physics by introducing hidden variables, we now reverse the direction of the argument and attempt to understand classical physics starting from quantum mechanics.

The problem It is a typical property of all classical-mechanical objects that they are never observed in the form of a superposition of two states. A classical object is for example always either at the location 1 or at the location 2, but never in a superposition of these two different spatial states. In terms of the two-slit experiment, this means that in experiments with classical objects, no interference can occur. The superposition which is possible for quantum objects, and thus the coherence of their states, is not allowed of classical objects. If we therefore attempt to describe classical objects by applying quantum mechanics to them, the problem arises as to how we can justify the decoherence of classical states without leaving the framework of quantum theory. We will try to derive the typically classical property that

the states of classical objects are not describable in terms of superpositions as an *emergent property*. The phase damping channel shows that entanglement with the environment in the form of scattering can play a role.

An example We imagine a ball or some other macroscopic body S^A. Its position states, described quantum mechanically, are denoted by $|0^A\rangle$ and $|1^A\rangle$. These can refer, for example, to the states behind the individual slits of a double-slit apparatus. In quantum-mechanical terms, the state

$$|\varphi^A\rangle = c_0|0^A\rangle + c_1|1^A\rangle \tag{15.18}$$

is possible; it would lead to an interference pattern. We have to take into account the fact that the ball is an open system. It interacts constantly with its environment, e. g. by scattering photons. Even in the dark, the cosmic background radiation remains. $|0^A\rangle$ and $|1^A\rangle$, as classical states, are unchanged by this scattering (stable states). We thus have a situation which is analogous to the scattering discussed in Sect. 15.1. If the ball is in the state $|0^A\rangle$, the photon numbered 1 will be scattered into the state $|0_1^E\rangle$; correspondingly for $|1^A\rangle$, the scattering leads to $|1_1^E\rangle$. As we have seen in Sect. 15.1, the state of the ball is described by the reduced density operator

$$\rho^A = \begin{pmatrix} |c_0|^2 & c_0c_1^*\langle 1_1^E|0_1^E\rangle \\ c_0^*c_1\langle 0_1^E|1_1^E\rangle & |c_1|^2 \end{pmatrix} . \tag{15.19}$$

In comparison to Eq. (15.9), we have taken $p = 0$, but have not required that the scattering states be orthogonal.

Not only one photon, but constantly a great number of photons are scattered. Correspondingly, the environment must be described by a product space $\mathcal{H}^E = \mathcal{H}_1^E \otimes \mathcal{H}_2^E \otimes \ldots$ containing many factor spaces. The numbers enumerate the two-dimensional factor spaces. Each factor space corresponds to a degree of freedom. The environment has many degrees of freedom. The entangled state $|\psi'^{AE}\rangle$ of the composite system S^{AE} after the scattering is

$$|\psi'^{AE}\rangle = c_1|0^A\rangle|0_1^E, 0_2^E \ldots\rangle + c_2|1^A\rangle|1_1^E, 1_2^E, \ldots\rangle . \tag{15.20}$$

The state of the ball is given by the reduced density operator

$$\rho'^A = \begin{pmatrix} |c_0|^2 & c_0c_1^*\langle 1_1^E|0_1^E\rangle\langle 1_2^E|0_2^E\rangle \ldots \\ c_0^*c_1\langle 0_1^E|1_1^E\rangle\langle 0_2^E|1_2^E\rangle \ldots & |c_1|^2 . \end{pmatrix} \tag{15.21}$$

The balls at different locations scatter photons in different ways; otherwise, we would not be able to distinguish them optically. The states $|0_j^E\rangle$ and $|1_j^E\rangle$, where $j = 1, 2, 3, \ldots$, are therefore all very different, and $|\langle 0_j^E|1_j^E\rangle| \ll 1$ holds. Since a large number of photons are scattered, a large number of these very small inner products occur in the non-diagonal elements of $\rho^{A'}$ in Eq. (15.21). Due to the entanglement with the scattered photons, the

state of the ball is transformed into a mixture which is no longer capable of exhibiting an interference pattern

$$\rho'^{A} \to |c_0|^2 |0^A\rangle\langle 0^A| + |c_1|^2 |1^A\rangle\langle 1^A| \,. \tag{15.22}$$

This density operator is only formally the same as that of a statistical mixture (blend) in which the ball is to be found *either* in the state $|0^A\rangle$ with the probability $|c_0|^2$ *or* in the state $|1^A\rangle$ with the probability $|c_1|^2$. All statistical statements about subsequent measurements are the same. An ignorance interpretation (cf. Sect. 4.3) is nevertheless not possible.

The superposition (15.20) continues to exist. It can, however, not be observed in the system S^A. Owing to the entanglement of S^A with its environment, the originally-present superposition has become non-local. The ball appears to be localised. It is remarkable that this is caused precisely by the fact that the state of the overall system is non-local. The dynamic process is based on a unitary transformation (compare Eq. (15.3)). It is hence reversible in principle (*recoherence*). If, however, one considers the many scattered photons, then it becomes clear that in practice, the process is irreversible. We shall return to the significance of these statements in connection with the problem of quantum measurements in Sect. 15.3.

What we have shown here for balls holds correspondingly for all systems S^A which (i) are subject to an entangling interaction with an environmental system S^E. This interaction is supposed (ii) to leave certain states of S^A unchanged (stable states) and (iii) to entangle each of them with many states of S^E which are very different (nearly orthogonal). The *environment-induced decoherence* which then occurs becomes stronger when the environmental system has more degrees of freedom. It transforms the state of S^A into a mixture of the stable states. These are the *classical states*, *since every superposition of these states – in the case that it ever appears at all – would decay very rapidly due to decoherence*, as we have seen above. The environmental system S^E can also consist of the internal degrees of freedom of a body. Internal degrees of freedom which are not explicitly taken into account in the description of the system S^A play the same role as an external environment. It is not necessary that the system S^E have macroscopic dimensions. It need only have a very large number of degrees of freedom.

The superposition of macroscopic quantum states We have seen that the decoherence approach can explain why superpositions of macroscopic states are so extraordinarily fragile. This does not however mean that they could not be realised in subtle experimental arrangements. A whole series of experiments demonstrate this. Here, we can only briefly list the corresponding topics: matter-wave interferometry with large molecules, Bose-Einstein condensates, quantum tunneling in SQUIDs, superconductors, photons in microwave cavities. Literature references can be found in Sect. 15.6. In these experiments, the role of environment-induced decoherence has also been studied. All the experiments can be explained by quantum mechanics without making any changes or additions to the theory. The macroscopic states (e. g. the 10^9 pairs of electrons in SQUIDs) were described as quantum states.

15.2.2 Schrödinger's Cat

The experiment Erwin Schrödinger [Sch 35] described the following thought experiment (cf. Fig. 15.3): a cat is closed in an opaque box together with a decaying radioactive source. Within one hour, a single radioactive decay may occur with the probability $\frac{1}{2}$. The decay causes a vial containing cyanide to be smashed and the cat to be killed. If one limits oneself to a description of the observations that an experimentalist would make in carrying out this experiment, there is no problem whatsoever: the trial is repeated a number of times with many cats and boxes. Each time, after one hour, the box is opened, or it is verified whether the cat is alive or dead. One finds that it is dead in half the trials and alive in the other half.

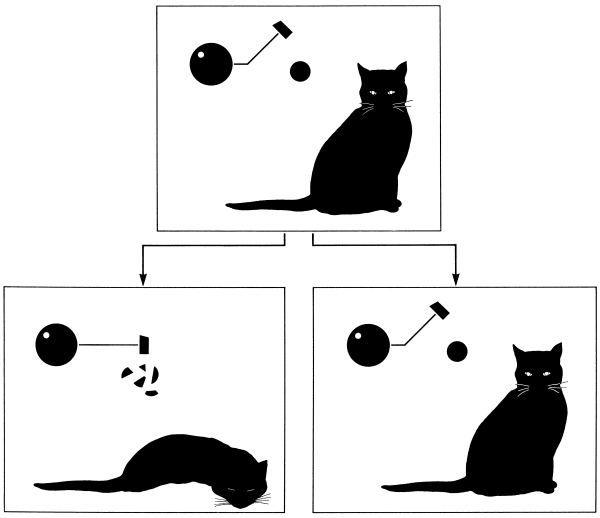


Figure 15.3: Upon opening the box, one finds Schrödinger's cat either alive or dead, with the probabilities 1/2. (From: Audretsch/Mainzer (eds.): *Wieviele Leben hat Schrödingers Katze?* ©Elsevier GmbH, Spektrum Akademischer Verlag, Heidelberg, 1990.)

Awakening dead cats to life If quantum mechanics is applied to the cat and a quantum state $|\text{alive}\rangle$ or $|\text{dead}\rangle$ is ascribed to it, there is at first also no problem. Before the measurement, i. e. before anyone opens the box and checks on its condition, the cat is represented by a superposition of the states "$|\text{alive}\rangle$" and "$|\text{dead}\rangle$" with a marking (see Fig. 15.4)

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|\text{alive}\rangle|\text{not decayed}\rangle + e^{i\phi}|\text{dead}\rangle|\text{decayed}\rangle) \; . \tag{15.23}$$

Such a superposition might seem unfamiliar in connection with cats. In any case, the experiment treated above, in which in the end a property "alive" or "dead" of the cat is measured, is described quite correctly.

There are, to be sure, plausible arguments based on experience as to why the state $|\psi\rangle$ of Eq. (15.23) cannot be realised as a stable state. It is essential that the state $|\psi\rangle$ of the cat and the

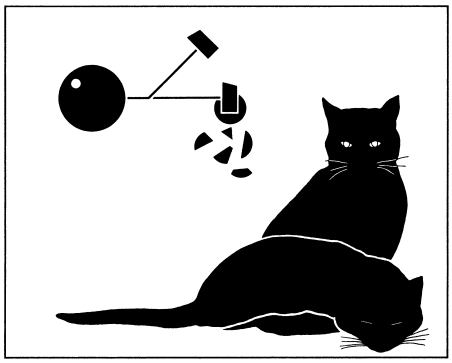


Figure 15.4: Is a superposition of the quantum states with a living and a dead cat present before opening the box? (From: Audretsch/Mainzer (eds.): *Wieviele Leben hat Schrödingers Katze?* ©Elsevier GmbH, Spektrum Akademischer Verlag, Heidelberg, 1990.)

radioactive source be in the form of a superposition. Despite the marking, superpositions can lead to interference, as we have seen for example in Sect. 8.7 in connection with which-path marking and quantum erasing. It is thus in principle not excluded that there could be another measurement outcome besides "alive" or "dead" (a "rotated basis"), in which an interference between a living and a dead cat would be observed. Such an interference has in any case never been seen, neither for macroscopic balls nor for cats.

Still worse; if as a result of the quantum-mechanical description of cats such states as $|\varphi\rangle = \frac{1}{\sqrt{2}}(|\text{alive}\rangle + |\text{dead}\rangle)$ were also to exist, then the following possibility would arise: one begins with an ensemble of cats in the state $|\text{dead}\rangle$ and carries out a projection measurement with the observable $|\varphi\rangle\langle\varphi| - |\varphi_\perp\rangle\langle\varphi_\perp|$, where $|\varphi_\perp\rangle = \frac{1}{\sqrt{2}}(|\text{alive}\rangle - |\text{dead}\rangle)$. This observable would then possibly also exist. Through a subsequent measurement of the observable $|\text{alive}\rangle\langle\text{alive}| - |\text{dead}\rangle\langle\text{dead}|$, in 50% of the cases the cats will be transferred to the state $|\text{alive}\rangle$. One could hence awaken dead cats to life.

In states involving cats, the implementation of a superposition is therefore not possible. *We have already seen in Sect. 15.2.1 that environment-induced decoherence (the system is in fact open) will prevent this. It transforms the superposition of Eq. (15.23) into a mixture of the stable states $|alive\rangle|not\ decayed\rangle$ and $|dead\rangle|decayed\rangle$.* Whether this already solves all the problems which occur in connection with the development of classical properties will be discussed in the next section.

15.3 The Quantum Measurement Process*

15.3.1 The Research Programme*

From its beginnings, quantum mechanics was constructed according to a dualistic dynamical scheme which we have likewise used in all the preceding sections: there are on the one

*The sections marked with an asterisk * can be skipped over in a first reading.

hand the quantum operations that describe the behaviour of the system between preparation and measurement, and on the other the quantum operations of the measurement itself. In the simplest case, these are unitary transformations or projections, respectively. Many physicists consider it dissatisfying to postulate two different dynamics. This has given rise to a research programme which aims at reducing the measurement dynamics, which was up to now introduced as a postulate, as far as possible or perhaps even completely to the unitary dynamics of the interaction between the system and the measurement apparatus. In this programme, the following requirements must be fulfilled:

(i) For various observables such as energy, spin, etc., there are different measuring devices with different interventions on the system. It must emerge from the dynamic evolution which observable is measured by a particular device.

(ii) The pointer of the measuring device must not move as a function of time after the measurement.

(iii) Pointers are classical systems. They must never be described as superpositions of different indicated values.

(iv) The computation must reflect the fact that in a single measurement, one and only one measurement outcome out of all the many possible ones is in fact indicated. This statement is also contained in the measurement postulates. However, no deterministic justification is required for precisely which value is indicated.

15.3.2 Pre-Measurement*

For simplicity, we treat the system S^A on which the measurement is to be performed as a qubit system with a quantum-mechanically described measuring device S^M, whose Hilbert space $\mathcal{H}^M$ likewise is of dimension two. The unitary measurement interaction on $\mathcal{H}^A \otimes \mathcal{H}^M$, which belongs to a measurement with the eigenstates $|0^A\rangle$ and $|1^A\rangle$ of S^A, is supposed to produce a marking (see Sect. 8.7). When $|0^A\rangle$ (or $|1^A\rangle$) is present, the measuring device is transferred into the state $|0^M\rangle$ (or $|1^M\rangle$). For the general state $|\varphi^A\rangle = c_0|0^A\rangle + c_1|1^A\rangle$ of S^A, this implies the following entanglement with the states of the measuring device:

$$\begin{aligned} & |\phi^{AM}\rangle = |\varphi^A\rangle|i^M\rangle = \left(c_0|0^A\rangle + c_1|1^A\rangle\right)|i^M\rangle \\ \rightarrow \quad & |\phi'^{AM}\rangle = c_0|0^A\rangle|0^M\rangle + c_1|1^A\rangle|1^M\rangle \end{aligned} \tag{15.24}$$

(compare the Stern-Gerlach experiment described in Sect. 13.2.2). $|i^M\rangle$ is an initial state of S^M. This dynamic evolution is often termed a *pre-measurement*.

The decomposition of the resulting state $|\phi'^{AM}\rangle$ in terms of a basis of $\mathcal{H}^A \otimes \mathcal{H}^M$ is not unique. For $c_0 = c_1 = \frac{1}{\sqrt{2}}$, the resulting state is a Bell state $|\Phi_+^{AM}\rangle = \frac{1}{\sqrt{2}}(|0^A\rangle|0^M\rangle + |1^A\rangle|1^M\rangle)$. We can also write it in the form

$$|\phi'^{AM}\rangle = |\Phi_+^{AM}\rangle = \frac{1}{\sqrt{2}}\left(|0_x^A\rangle|0_x^M\rangle + |1_x^A\rangle|1_x^M\rangle\right) \ . \tag{15.25}$$

If $|\phi'^{AM}\rangle$ is supposed to be the state for which the measuring device has measured the observable

$$Z^A = z_0|0^A\rangle\langle 0^A| + z_1|1^A\rangle\langle 1^A| \ , \tag{15.26}$$

then one can just as well consider the apparatus to be a device for measuring the observable $X^A = x_0|0_x^A\rangle\langle 0_x^A| + x_1|1_x^A\rangle\langle 1_x^A|$. It is thus not yet determined which observable will in the end be measured in the unitary evolution (15.24). Or, from another point of view: if $|0^M\rangle$ and $|1^M\rangle$ are the classical indicator states, then it is not excluded that the result of the pre-measurement is a marking with the superpositions of classical pointer states. Requirement (i) has not yet been fulfilled. With a view to Sect. 15.2.1, the assumption suggests itself that this problem can be solved by taking the environment into account.

The other requirements are also not fulfilled by the pre-measurement. We can demonstrate this using the example of requirement (ii). Unitary evolutions are reversible. On the other hand, the measurement process is irreversible. We will investigate this with the example of the simple unitary evolution $U = e^{-iHt}$ where the Hamiltonian is

$$H^{AM} = g\,\sigma_z^A \otimes \sigma_y^M \tag{15.27}$$

on $\mathcal{H}^A \otimes \mathcal{H}^M$. Let the initial state be the product state

$$|\phi^{AM}(t=0)\rangle = \left(c_0|0^A\rangle + c_1|1^A\rangle\right)|0_x^M\rangle \tag{15.28}$$

with $|c_0|^2 + |c_1|^2 = 1$. An auxiliary calculation which we shall not reproduce here (cf. Problem 15.1) leads to

$$\begin{aligned} |\phi^{AM}(t)\rangle &= c_0|0^A\rangle\left\{\sin\left(\frac{\pi}{4}+gt\right)|0^M\rangle + \cos\left(\frac{\pi}{4}+gt\right)|1^M\rangle\right\} \\ &+ c_1|1^A\rangle\left\{\sin\left(\frac{\pi}{4}-gt\right)|0^M\rangle + \cos\left(\frac{\pi}{4}-gt\right)|1^M\rangle\right\}\ . \end{aligned} \tag{15.29}$$

We obtain a time-dependent entanglement which leads at the time $t = \frac{\pi}{4g}$ to the desired state $|\Phi'^{AM}\rangle$ of Eq. (15.24). At the time $t = \frac{2\pi}{g}$, however, the non-entangled initial state $|\Phi^{AM}(t=0)\rangle$ is again obtained. The pre-measurement hence does not lead to a stable marking of the composite system.

15.3.3 Entanglement with the Environment Fixes the Observable*

A first step towards a solution of these problems consists of extending the systems S^A and S^M to include the environment S^E, which in a first step is assumed to consist of a single qubit system. The state of the composite system S^{AME} lies within $\mathcal{H}^A \otimes \mathcal{H}^M \otimes \mathcal{H}^E$. We couple the measuring device S^M to the environment S^E via a dynamic with the Hamiltonian H^{ME}, which acts only on $\mathcal{H}^M \otimes \mathcal{H}^E$ and thus leaves the state of S^A obtained from the pre-measurement unchanged

$$\mathbb{1}^A \otimes H^{ME} = g\,\mathbb{1}^A \otimes \sigma_z^M \otimes \sigma_z^E\ . \tag{15.30}$$

The states $|0^E\rangle$ and $|1^E\rangle$ form an ONB of $\mathcal{H}^E$.

The initial state at the time $t = 0$ following the pre-measurement is, with $|\phi'^{AM}\rangle$ from Eq. (15.24),

$$|\psi(t=0)\rangle = \left\{c_0|0^A\rangle|0^M\rangle + c_1|1^A\rangle|1^M\rangle\right\}\left(\alpha|0^E\rangle + \beta|1^E\rangle\right)\ . \tag{15.31}$$

The unitary evolution caused subsequently by $\mathbb{1}^A \otimes H^{ME}$ leads to a time-dependent entanglement with the environment (compare Problem 15.2):

$$|\psi(t)\rangle = c_0|0^A\rangle|0^M\rangle|\omega_0^E(t)\rangle + c_1|1^A\rangle|1^M\rangle|\omega_1^E(t)\rangle \tag{15.32}$$

where

$$|\omega_0^E(t)\rangle = \alpha\exp(-igt)|0^E\rangle + \beta\exp(igt)|1^E\rangle = |\omega_1^E(-t)\rangle\ . \tag{15.33}$$

Is it still undetermined which states are occupied by the system S^A and the measuring device S^M? The rearrangement of Eq. (15.24) into Eq. (15.25) for $c_0 = c_1 = \frac{1}{\sqrt{2}}$ is possible only when the summands contain no differing factors. The states $|\omega_0^E\rangle$ and $|\omega_1^E\rangle$ in Eq. (15.32) are such differing "factors". The entanglement with the environment, consisting of only a single qubit, has already produced the effect that only a correlation between the eigenstates of σ_z^A and σ_z^M, but no longer between those of σ_x^A and σ_x^M, exists[1]. The apparatus with the measurement dynamics H^{ME} is therefore associated with the observable Z^A. Requirement (i) is fulfilled. However, only a unitary transformation has been carried out. The states are still periodically entangled and disentangled. At the time $t = \frac{2\pi}{g}$, the state $|\psi(t=0)\rangle$ is again obtained.

15.3.4 Entanglement with Many Degrees of Freedom of the Environment*

Collapse and revival Realistic environments have a large number of degrees of freedom. We are thus dealing with a composite system in a Hilbert space $\mathcal{H} = \mathcal{H}^A \otimes \mathcal{H}^M \otimes \mathcal{H}^E_{(1)} \otimes \ldots \otimes \mathcal{H}^E_{(N)}$, where N is a very large number. We discuss the simple case that all the Hilbert spaces are two-dimensional, and proceed again as in the previous section.

Before the interaction between the measuring device and the environment is switched on, the state

$$|\psi(t=0)\rangle = \left\{c_0|0^A\rangle|0^M\rangle + c_1|1^A\rangle|1^M\rangle\right\} \otimes \left\{\prod_{k=1}^{N}\left(\alpha_k|0^E_{(k)}\rangle + \beta_k|1^E_{(k)}\rangle\right)\right\} \tag{15.34}$$

with $|\alpha_k|^2 + |\beta_k|^2 = 1$ is present. The interaction is supposed to be given by the simple Hamiltonian

$$\mathbb{1}^A \otimes H^{ME} = \sum_{k=1}^{N} \mathbb{1}^A \otimes H_k^{ME} \tag{15.35}$$

[1] Making use of the following theorem, the proof can be made more precise [Bub 97, Chap. 5.5]: A representation of a state $|\psi\rangle \in \mathcal{H} \otimes \mathcal{H}' \otimes \mathcal{H}''$ as a *tri-orthogonal decomposition*

$$|\psi\rangle = \sum_i c_i|u_i\rangle|v_i\rangle|w_i\rangle$$

with basis vectors $\{|u_i\rangle\} \in \mathcal{H}$, $\{|v_i\rangle\} \in \mathcal{H}'$ and $\{|w_i\rangle\} \in \mathcal{H}''$ does not always exist. But if it exists, it is unique.

with

$$\mathbb{1}^A \otimes H_k^{ME} := g_k \mathbb{1}^A \otimes \sigma_z^M \otimes \sigma_{z(k)}^E \otimes \prod_{j \neq k} \mathbb{1}_{(j)}^E \ . \tag{15.36}$$

The state $|\psi(t)\rangle$ of the composite system at the time t is obtained via the unitary transformation belonging to $\mathbb{1}^A \otimes H^{ME}$. After some computation (see Problem 15.2), for which we use the result (15.32), we find

$$|\psi(t)\rangle = c_0 |0^A\rangle |0^M\rangle |\Omega_0^E(t)\rangle + c_1 |1^A\rangle |1^M\rangle |\Omega_1^E(t)\rangle \tag{15.37}$$

where

$$|\Omega_0^E(t)\rangle := \prod_{k=1}^{N} (\alpha_k \exp(-i g_k t) |0_{(k)}^E\rangle + \beta_k \exp(i g_k t) |1_{(k)}^E\rangle) =: |\Omega_1^E(-t)\rangle \ . \tag{15.38}$$

The environment states are normalised: $\langle \Omega_0^E(t) | \Omega_0^E(t) \rangle = 1, \quad \langle \Omega_1^E(t) | \Omega_1^E(t) \rangle = 1$. But in general, they are not orthogonal at all times,

$$r(t) := \langle \Omega_0^E(t) | \Omega_1^E(t) \rangle \ . \tag{15.39}$$

For the description of the effects of the measurement on the system S^A due to the measuring device S^M, we must employ the reduced density operator of the system S^{AM}, taking the trace over the degrees of freedom of the environment:

$$\begin{aligned} \rho^{AM} = \mathrm{tr}_E[|\psi(t)\rangle\langle\psi(t)|] \quad &= \quad |c_0|^2 |0^A\rangle\langle 0^A| \otimes |0^M\rangle\langle 0^M| \\ &+ \quad |c_1|^2 |1^A\rangle\langle 1^A| \otimes |1^M\rangle\langle 1^M| \\ &+ \quad r(t) c_0 c_1^* |0^A\rangle\langle 1^A| \otimes |0^M\rangle\langle 1^M| \\ &+ \quad r^*(t) c_0^* c_1 |1^A\rangle\langle 0^A| \otimes |1^M\rangle\langle 0^M| \ . \end{aligned} \tag{15.40}$$

One readily finds (compare Problem 15.2) the explicit time dependence of $r(t)$:

$$r(t) = \prod_{k=1}^{N} \{\cos 2g_k t + i(|\alpha_k|^2 - |\beta_k|^2) \sin 2g_k t\} \ . \tag{15.41}$$

By construction, $r(t=0) = 1$ and $|r(t)|^2 \leq 1$.

The behaviour of $r(t)$ for long times is important in view of the requirement (ii). As is shown by Eq. (15.41), $r(t)$ is composed of periodic functions with many different frequencies $2g_k$. As is known from statistical mechanics, from quantum optics and from other areas of theoretical physics, such functions exhibit *collapse and revival*. Beginning at $t = 0$ with $r = 1$, $|r(t)|$ initially decreases and approaches the value zero, but after a long time it can again increase, again fall to zero, and so on. $|r(t)|$ can return arbitrarily closely[2] to $|r(t)| = 1$. With increasing N, this *recoherence* is however shifted to later and later times (see Fig. 15.5).

[2]Revival or recurrence can be interpreted more precisely: let $\mathcal{H}$ be a finite-dimensional Hilbert space and U a unitary operator. Then for every $\varepsilon > 0$, there is a natural number q such that $\sup\{\|U^q - \mathbb{1}|\varphi\rangle\| \ : \ |\varphi\rangle \in \mathcal{H}, \|\varphi\| = 1\} < \varepsilon$.
If the application of U is repeated often enough, one will return arbitrarily closely to the initial state [SSS 04].

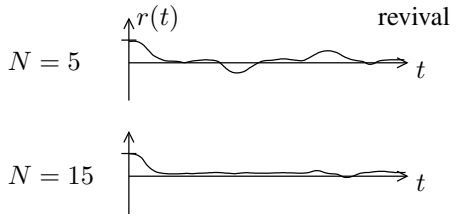


Figure 15.5: Collapse and revival (for a particular choice of parameters).

$r(t)$ rapidly attains a value which is proportional to 2^{-N} (cf. Sect. 15.6). As shown by Eq. (15.40), this corresponds to a rapid decoherence process in the bases $\{|0^A\rangle, |1^A\rangle\}$ and $\{|0^M\rangle, |1^M\rangle\}$. S^{AM} goes from a pure state to a mixture ρ^{AM}. In particular, the marker states $|0^M\rangle$ and $|1^M\rangle$ can no longer interfere with each other. They have become classical states which are called *pointer states*. Hence, requirement (iii) is also fulfilled.

The decoherence process induced by the environment S^E has marked out the basis of the pointer states in the reduced density operator ρ^{AM}. The result for the subsystem S^{AM} is the reduced density operator

$$\rho^{AM} = |c_0|^2 |0^A, 0^M\rangle\langle 0^A, 0^M| + |c_1|^2 |1^A, 1^M\rangle\langle 1^A, 1^M| \; . \tag{15.42}$$

A statistical mixture of the states $|0^A, 0^M\rangle$ and $|1^A, 1^M\rangle$ from $\mathcal{H}^A \otimes \mathcal{H}^M$ with the probabilities $p_0 = |c_0|^2$ and $p_1 = |c_1|^2$ is likewise described by the density operator ρ^{AM}. Such a statistical mixture corresponds according to the measurement postulate precisely to the result of a non-selective measurement of the observables Z^A. Are then all the requirements from Sect. 15.3.1 now fulfilled?

Which pointer states? With this restricted point of view which refers only to S^{AM} and does not consider the environment S^E, the argumentation is however still incomplete, since we have once again acquired a problem with requirement (i). Thus ρ^{AM} is the reduced density operator of an entangled subsystem and therefore not the result of a preparation process which leads to a statistical mixture. *An ignorance interpretation (see Sect. 4.3) is not possible*. The composite system is in a pure state. The density operator ρ^{AM} has arbitrarily many ensemble decompositions, for example also

$$\rho^{AM} = p_u |u^A, u^M\rangle\langle u^A, u^M| + p_v |v^A, v^M\rangle\langle v^A, v^M| \; . \tag{15.43}$$

Has therefore possibly a quite different observable $U^A = u|u^A\rangle\langle u^A| + v|v^A\rangle\langle v^A|$ with the pointer states $|u^M\rangle$ and $|v^M\rangle$ been measured?

We must not leave out the environment. Interactions with the environment are constantly present. If, as above, they take the form

$$H^{ME} = Z^M \otimes H^E \tag{15.44}$$

with an Hermitian operator H^E, then H^{ME} commutes with the *pointer observable* Z^M:

$$[H^{ME}, Z^M]_- = 0 \; . \tag{15.45}$$

The states $|0^M\rangle$ and $|1^M\rangle$, and thereby also the correlations with the states $|0^A\rangle$ and $|1^A\rangle$, remain unchanged under the influence of the environment. The stability of these correlations distinguishes the pointer basis $\{|0^M\rangle, |1^M\rangle\}$ from other bases.

The diagonalisation of ρ^{AB} by environment-induced decoherence and the stability of the correlations between S^A and S^M under the continuing influence of the environment leads to fixing of the time-invariant classical pointer states. Z^A, as a measured observable, is likewise fixed. Thus, the requirements (i) through (iii) of a dynamically-founded theory of the measurement process for non-selective (!) measurements are essentially fulfilled for this very simple model through coupling to the environment. The entanglement with the environment S^E is, however, not destroyed, as is shown by the possibility of "recurrence" in the far distant future. ρ^{AM} does not describe a statistical mixture. An ignorance interpretation (see Sect. 4.3) is not possible. The decision as to whether this should be regarded as a serious defect of the model is left to the reader.

15.4 Has the Problem of Measurements been Solved?*

An additional serious defect is in any case present. *A measurement leads to a measurement outcome.* This is its most important characteristic. The requirement (iv) in Sect. 15.3.1 is not fulfilled. Environment-induced decoherence cannot explain why in every individual experiment, always only one of the many possible indicator results of a measuring device is actually obtained. It has not been shown that such an event occurs. This is, however, precisely the most elementary experience of an experimenter who carries out measurements. It is a part of the measurement postulates. Decoherence is necessary but not sufficient for the explanation of the measurement process. *The measurement problem is thus (in the standard interpretation) unsolved.*

The gap in our chain of explanation refers typically to an individual process and thus as a consequence to selective measurements in general, since they are based upon individual measurements. With the state ρ^{AM} of the reduced density operator in the previous section, we obtained a mixture. If we interpret this as the result of a non-selective measurement and hence as a statistical mixture, we have already assumed in advance, in an inacceptable manner, that requirement (iv) could be fulfilled.

It seems to be the case that one cannot justify dynamically the requirement (iv) within the framework of the existing quantum theory including the standard interpretation. If this is true, we have two possibilities. We can change the quantum theory by for example postulating other dynamic equations (see Sect. 15.6). The second possibility consists of leaving the theory unchanged in its present mathematical formulation but of finding a different interpretation for it. We shall give a brief example of this.

15.5 The Many-Worlds Interpretation*

The *many-worlds interpretation*, which is also called the *Everett interpretation*, includes the observer as a system within the composite system [Eve 57]. He corresponds formally to the

*The sections marked with an asterisk * can be skipped over in a first reading.

system S^M as in Sect. 15.3.2. In the following, we will understand S^M in this sense. As in the pre-measurement, the state of S^{AM} evolves, if S^A is a qubit, into the state (15.25). The states $|0^M\rangle$ and $|1^M\rangle$ in this connection are *memory states* of the observer. This already describes the measurement process. The state $|\phi^{AM}\rangle$ is not transformed subsequently into one of the states (e. g. $|0^A\rangle|0^M\rangle$), but rather both *branches* $|0^A\rangle|0^M\rangle$ and $|1^A\rangle|1^M\rangle$ are separately realised. Their realisation manifests itself however in this interpretation in different real worlds which do not interact with one another. Since the states $|0^M\rangle$ and $|1^M\rangle$ are supposed to be memory states, an observer with a well-determined memory, for example $|1^M\rangle$, can recall only the measurement outcome belonging to $|1^A\rangle$ and he correspondingly finds the system S^A also in the state $|1^A\rangle$ and not in the state $|0^A\rangle$. Predictions about what will be observed in a subsequent experiment are probability statements. In this subsequent experiment as well, again everything which is potentially possible will be realised. The world splits again and again with each measurement. Reversing the quote from Niels Bohr in Sect. 2.5.1, one could say in summary: "There is no classical world".

It is the essence of an interpretation that it cannot be empirically refuted. This also holds for the many-worlds interpretation, which we have sketched here only superficially. Furthermore, everyone is free to choose that interpretation which he or she holds to be the most suitable for sound (metaphysical) reasons. We analysed this point in Sect. 2.5. The many-worlds interpretation is bizarre, but it has, in addition to many uncertainties (compare Sect. 15.6) and details which are not yet treated in a satisfactory fashion, also some advantages: requirement (iv) for the theory of the measurement process is fulfilled. Furthermore, the observer is included within the system. "Everything" is described. There is only a single closed system with a state vector which evolves in a strictly deterministic manner and which represents the entire cosmological universe. This could possibly facilitate our understanding of the quantum physics of the extremely early stages of the universe. On the other hand, we can identify various problems. For example: as in the pre-measurement of Sect. 15.3.2, one must ask why the splitting-up into different worlds does not take place according to Eq. (15.25). Will the observer S^M recall a measurement of the observable Z^A or that of the observable X^A? It appears reasonable that we once again consider decoherence due to the environment for the solution of this problem (which is problem (i) from Sect. 15.3.1).

In the next and final chapter, we shall return to "hard" theoretical material and present some proofs which were thus far missing.

15.6 Complementary Topics and Further Reading

- Review articles on decoherence: [Joo 96], [Zeh 96], [Bub 97, Chap. 5.4], [Zeh 00], [Joo 06], [PZ 02], [Leg 02], [Zur 02], [Zur 03], [Sch 04].
- Many aspects of the topic of decoherence are treated in the anthology [GJK 96].
- A comparison of decoherence rates and relaxation rates as well as time scales for decoherence: [Joo 96], [PZ 02].
- The time dependence for the vanishing of decoherence between spatially-separated components of a wavefunction (localisation rates): [Joo 96].

- A series of very readable essays on Schrödinger's cat, on the relationship between microphysics and macrophysics, and on the formation of the classical world can be found in [AM 96].
- We have studied only a very simplified model of environment-induced decoherence. One should be cautious in generalising the results: [PZ 02].
- Interferometry with macromolecules and decoherence: [ANZ 02], [HUH 03].
- The superposition of macroscopic systems in a SQUID: [FPC 00], [VdW 00].
- Experiments on entanglement, decoherence, and cat-states in cavity resonators: [HRB 02], [RBH 01].
- Experiments on the superposition of macroscopic quantum states; review articles: [Leg 02], [Sch 05].
- Classic articles on the role of decoherence in the measurement process: [Zeh 70], [Zur 81], [Zur 82].
- Review articles and books on the theory of the quantum measurement process: [Zur 82], [Zur 91], [BLM 91], [Bub 97, Chap. 8], [Mit 98], [Aul 00, Chap. IV], [PZ 02], [Leg 02].
- The critical letters which were published in answer to the article [Zur 91] are also worth reading: Physics Today **44** (4), 13-15 and 81-90, April 1993.
- The *vacuum* of Minkowski space from the point of view of an accelerated observer has the structure of an entangled state of quantum systems which are at different locations within the flat space-time. Measurements with accelerated 2-level detectors can yield unexpected non-local effects: [AM 94a], [AM 94b]. Review article: [PT 04].
- Review articles and books on the many-world interpretation: [DG 73], [Deu 96], [Bar 00], [Vai 01].
- Brief accounts of the many-world interpretation: [Pri 81, Chap. 3.6], [d'Es 95, Chap. 12], [Bub 97, Chap. 8.2], [Hom 97, Chap. 2] [Mit 98, Chap. 3.2], [Aul 00, Chap. 15], [Lal 01].
- Quantum theory has been related to *consciousness* by numerous authors in many different ways. As early as 1939, the hypothesis was proposed that the activity of the observer's consciousness might be responsible for the fact that in the measurement process, a transition takes place into a particular outgoing state (the *London-Bauer interpretation* [LB 39]). The standard theory of quantum mechanics including the postulates for the measurement process is not modified in this approach. The boundary between the measuring device and the system on which the measurement is performed is shifted into the brain of the observer. This representation is thus not a modification of quantum theory, but rather another interpretation. For a detailed treatment, see [Shi 63].
- The attempts to understand human consciousness itself on the basis of quantum theory have a different intention. An example as well as references to more detailed literature can be found in [Sta 93].

- *Collapse theories* have as their goal the description of the dynamic evolution between measurements and within the measurements themselves by means of a single dynamic equation. To this end, the Schrödinger equation is modified by including stochastic and nonlinear terms. The standard model in this case contains phenomenological parameters ([GRW 86]). Review articles on the description of the measurement process with modified Schrödinger and von Neumann equations (state reduction as a dynamic process): see [Sta 96], [Lal 01].

- If one wants to describe the early state of the whole universe as a quantum state, then the concept of reducing coherence due to an interaction with the environment fails, since the universe by definition has no environment. An attempt to solve this problem is represented by the theory of *consistent histories*. It resembles the "sum over histories" formulation of quantum theory given by Feynman. The question as to which histories are consistent is investigated. Various probabilities are attributed to different histories. An introduction can be found in [Gri 02].

- In this book, we have described the further development of quantum theory in recent years. The concepts of quantum information theory played an important role. This makes it interesting to try to formulate the fundamentals of quantum theory with reference to the manipulation of information. In [CBH 03] and [Bub 05a], an information-theoretical approach is described. It is based on three no-go postulates. One of these is e. g. the no-signaling requirement: no local measurements on a system should have an instantaneous influence on the state (reduced density matrix) of another system.

15.7 Problems for Chapter 15

Prob. 15.1 [for 15.1]: Calculate the effect of the phase damping channel on the density operator

$$\rho = \frac{1}{2}(\mathbb{1} + \mathbf{r}\boldsymbol{\sigma}) \; . \tag{15.46}$$

How does the Bloch vector $\mathbf{r}$ change?

Prob. 15.2 [for 15.3]: Derive Eqs. (15.29), (15.32), and (15.37). Begin by computing the dyadic decomposition of the operators $\exp(\sigma_z^A \otimes \sigma_y^M)$ and $\exp(\sigma_z^A \otimes \sigma_z^M)$ in suitable bases of $\mathcal{H}^A \otimes \mathcal{H}^M$.

Prob. 15.3 [for 15.3]: Derive Eq. (15.40).

© 2007 WILEY-VCH Verlag GmbH & Co.

16 Two Implementations of Quantum Operations*

We want to append the proof of Kraus' theorem from Sect. 14.1.2. In addition, the implementation of quantum operations asserted in that section will be demonstrated here. Based upon this, the implementation of a general quantum measurement can be described. All these procedures are based on entanglement with an ancilla system. We thus pursue the investigations from Sections 7.7 and 7.8 further here.

16.1 The Operator-Sum Decomposition*

Relative states and index states We consider the bipartite system S^{AB} with the Hilbert space $\mathcal{H}^A \otimes \mathcal{H}^B$ and take for simplicity $\dim \mathcal{H}^A = \dim \mathcal{H}^B = d$. In $\mathcal{H}^A$ and $\mathcal{H}^B$, we introduce the ONB $\{|a_n^A\rangle\}$ and $\{|e_n^B\rangle\}$ and construct with them a maximally-entangled, non-normalised state:

$$|\tilde{\psi}^{AB}\rangle = \sum_{n=1}^{d} |a_n^A, e_n^B\rangle \tag{16.1}$$

for which

$$|a_n^A\rangle = \langle e_n^B|\tilde{\psi}^{AB}\rangle \tag{16.2}$$

holds. We can consider $|e_n^B\rangle$ to be a marker state for $|a_n^A\rangle$.

This idea can be generalised. We can obtain an arbitrary state $|\phi^A\rangle$

$$|\phi^A\rangle = \sum_n c_n |a_n^A\rangle \tag{16.3}$$

of $\mathcal{H}^A$ by projection in $\mathcal{H}^B$. To do so, we generate in $\mathcal{H}^B$ the *index state* belonging to $|\phi^A\rangle$

$$|\phi^{*B}\rangle := \sum_n c_n^* |e_n^B\rangle \ . \tag{16.4}$$

We then find with Eq. (16.1)

$$|\phi^A\rangle = \langle \phi^{*B}|\tilde{\psi}^{AB}\rangle \tag{16.5}$$

*The chapters marked with an asterisk * can be skipped over in a first reading.

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

and likewise

$$|\phi^{*B}\rangle = \langle\phi^A|\tilde{\psi}^{AB}\rangle \,. \tag{16.6}$$

$|\phi^A\rangle$ is called the *relative state*. Via the maximally-entangled ancilla state $|\tilde{\psi}^{AB}\rangle$, which depends only on the two ONBs, a simple relation is obtained between an arbitrary state $|\phi^A\rangle \in \mathcal{H}^A$ and its index state $|\phi^{*B}\rangle$ in $\mathcal{H}^B$. $|\tilde{\psi}^{AB}\rangle$ establishes an unambiguous mapping

$$|\phi^A\rangle \leftrightarrow |\phi^{*B}\rangle \,, \tag{16.7}$$

which is *conjugate linear*. We have

$$|\phi^A\rangle = a_1|\phi_1^A\rangle + a_2|\phi_2^A\rangle, \qquad |\phi^{*B}\rangle = a_1^*|\phi_1^{*B}\rangle + a_2^*|\phi_2^{*B}\rangle \,. \tag{16.8}$$

The proof is carried out by expansion in terms of the basis vectors.

Proof of Kraus' theorem We want to prove the following statement, which was already asserted in Sect. 14.1.2: *A mapping $\rho \to \tilde{\rho}' = \mathcal{E}(\rho)$ is a quantum operation if and only if an operator-sum decomposition*

$$\mathcal{E}(\rho) = \sum_i K_i \rho K_i^\dagger \tag{16.9}$$

with linear operators K_i exists, which fulfill the condition

$$\sum_i K_i^\dagger K_i \leq \mathbb{1} \tag{16.10}$$

and which map the incoming Hilbert space onto the outgoing Hilbert space. The condition (16.10) is equivalent to the statement that the trace does not increase:

$$\mathrm{tr}[\mathcal{E}(\rho)] \leq \mathbb{1} \tag{16.11}$$

($\mathrm{tr}[\rho] = 1$). $\tilde{\rho}'$ is again a density operator.

For the proof in the one direction, we construct the superoperator $\mathcal{E}$ according to Eq. (16.9) with the linear operators K_i. We show that $\mathcal{E}^A \otimes \mathbb{1}^B$ transforms an arbitrary positive operator π^{AB} from $\mathcal{H}^A \otimes \mathcal{H}^B$ into a positive operator. Let $|\psi^{AB}\rangle$ be an arbitrary vector in $\mathcal{H}^A \otimes \mathcal{H}^B$. We choose an index i and form

$$|\phi_i^{AB}\rangle := (K_i^{A\dagger} \otimes \mathbb{1}^B)|\psi^{AB}\rangle \,. \tag{16.12}$$

Then $K_i^A \otimes \mathbb{1}^B$ is a positive operator on $\mathcal{H}^A \otimes \mathcal{H}^B$ owing to

$$\langle\psi^{AB}|(K_i^A \otimes \mathbb{1}^B)\pi^{AB}(K_i^{A\dagger} \otimes \mathbb{1}^B)|\psi^{AB}\rangle = \langle\phi_i^{AB}|\pi^{AB}|\phi_i^{AB}\rangle \geq 0 \,. \tag{16.13}$$

This holds also for the sum over i. The superoperator $\mathcal{E}^A$ given in the form of Eq. (16.9) is not only positive, but also completely positive.

For the proof in the converse direction, we begin with the maximally-entangled state $|\tilde{\psi}^{AB}\rangle\langle\tilde{\psi}^{AB}|$ of Eq. (16.1). Since $\mathcal{E}^A$ is by assumption a quantum operation, the application of $\mathcal{E}^A \otimes \mathbb{1}^B$

$$(\mathcal{E}^A \otimes \mathbb{1}^B)|\tilde{\psi}^{AB}\rangle\langle\tilde{\psi}^{AB}| =: C^{AB} \tag{16.14}$$

yields a positive operator C^{AB}.

In the second step of the proof, we write out C^{AB} with Eq. (16.1) and making use of the linearity of $\mathcal{E}^A$

$$\begin{aligned} C^{AB} &= \left(\mathcal{E}^A \otimes \mathbb{1}^B\right) \left(\sum_n |a_n^A, e_n^B\rangle\right) \left(\sum_{n'} \langle a_{n'}^A, e_{n'}^B|\right) \\ &= \sum_{n,n'} \mathcal{E}^A\left(|a_n^A\rangle\langle a_{n'}^A|\right) \otimes \left(|e_n^B\rangle\langle e_{n'}^B|\right) \,. \end{aligned} \tag{16.15}$$

We take with Eqs. (16.3) and (16.4) the expectation value with the index state $|\phi^{*B}\rangle$:

$$\langle\phi^{*B}|C^{AB}|\phi^{*B}\rangle = \sum_{n,n'} c_n c_{n'}^* \mathcal{E}^A(|a_n^A\rangle\langle a_{n'}^A|) = \mathcal{E}^A(|\phi^A\rangle\langle\phi^A|) \,. \tag{16.16}$$

Here, we have again made use of the linearity of $\mathcal{E}^A$. Equation (16.16) demonstrates that the complete information about the action of the superoperators $\mathcal{E}^A$ on the basis of the Liouville space $\mathbb{L}^A$ is contained in the operator C^{AB}.

In the third step, we choose an ensemble decomposition of C^{AB},

$$C^{AB} = \sum_i |\tilde{c}_i^{AB}\rangle\langle\tilde{c}_i^{AB}| \tag{16.17}$$

and rearrange the left-hand side of Eq. (16.16)

$$\mathcal{E}^A(|\phi^A\rangle\langle\phi^A|) = \sum_i \langle\phi^{*B}|\tilde{c}_i^{AB}\rangle\langle\tilde{c}_i^{AB}|\phi^{*B}\rangle \,. \tag{16.18}$$

We introduce the operators K_i^A defined by their action on $|\phi^A\rangle$:

$$K_i^A|\phi^A\rangle := \langle\phi^{*B}|\tilde{c}_i^{AB}\rangle \tag{16.19}$$

and obtain from Eq. (16.18)

$$\mathcal{E}^A(|\phi^A\rangle\langle\phi^A|) = \sum_i K_i^A|\phi^A\rangle\langle\phi^A|K_i^{A\dagger} \,. \tag{16.20}$$

Due to the linearity of $\mathcal{E}^A$, it then holds for all density operators ρ^A on $\mathcal{H}^A$ that:

$$\mathcal{E}^A(\rho^A) = \sum_i K_i^A \rho^A K_i^{A\dagger} \tag{16.21}$$

since the mapping $|\phi^A\rangle \leftrightarrow |\phi^{*B}\rangle$ is conjugate linear. With Eq. (16.19), the operators K_i^A are also linear operators. Furthermore, since the quantum operation $\mathcal{E}^A(\rho^A)$ of Eq. (16.21) does not conserve the trace for all ρ^A, we obtain the condition (16.10). This proves the theorem on operator-sum decomposition in both directions.

The choice of the two orthonormal bases $\{|a_n^A\rangle\}$ and $\{|e_n^B\rangle\}$, like the ensemble decomposition (16.17), was arbitrary. This once more demonstrates that the operation elements K_i^A for a given superoperator $\mathcal{E}^A$ are not uniquely determined.

16.2 The Unitary Implementation of Quantum Operations*

Trace-conserving quantum operations A *trace-conserving quantum operation* $\mathcal{E}^A$ on the system S^A

$$\mathrm{tr}[\mathcal{E}^A(\rho^A)] = \mathrm{tr}[\rho^A] = 1 \tag{16.22}$$

is also called a *complete quantum operation*. The associated decomposition operators obey the completeness relation

$$\sum_i K_i^{A\dagger} K_i^A = \mathbb{1} \ . \tag{16.23}$$

For a unitary implementation of the quantum operation, we again extend the system S^A by adding an ancilla system S^B. The dimension of $\mathcal{H}^B$ is assumed to agree with the number of decomposition operators.

In $\mathcal{H}^B$, we choose a basis $\{|i^B\rangle\}$. Let S^B be initially in the state $|0^B\rangle$ and the composite system in the product state $|\psi^A\rangle|0^B\rangle$. This notation is different from that used in Sect. 14.1.2. Using the K_i^A, we define an operator $\hat{U}^{AB}$ through its action on states $|\psi^A\rangle|0^B\rangle$:

$$\hat{U}^{AB}|\psi^A\rangle|0^B\rangle := \sum_i K_i^A|\psi^A\rangle|i^B\rangle \ . \tag{16.24}$$

For arbitrary states $|\psi^A\rangle$ and $|\phi^A\rangle$, we then have with the completeness relation (16.23)

$$\begin{aligned} \langle\psi^A|\langle 0^B|\hat{U}^{AB\dagger}\hat{U}^{AB}|\phi^A\rangle|0^B\rangle &= \sum_i \langle\psi^A, 0^B|K_i^{A\dagger}K_i^A|\phi^A, 0^B\rangle \\ &= \langle\psi^A, 0^B|\phi^A, 0^B\rangle \ . \end{aligned} \tag{16.25}$$

$\hat{U}^{AB}$ conserves the inner products on states of the form $|\psi^A, 0^B\rangle$. With the corollary obtained in Sect. 13.3.5, we can then extend $\hat{U}^{AB}$ to a unitary operator U^{AB} which acts on the whole of $\mathcal{H}^A \otimes \mathcal{H}^B$. We assume that such operators have a physical implementation.

*The sections marked with an asterisk * can be skipped over in a first reading.

We make use of Eq. (16.24). The subsystem S^A is transformed into the state

$$\begin{aligned}\mathrm{tr}_B[U^{AB}(|\psi^A\rangle\langle\psi^A|\otimes|0^B\rangle\langle 0^B|)U^{AB\dagger}] \\ &= \mathrm{tr}_B[\sum_{i,i'} K_i^A|\psi^A\rangle\langle\psi^A|K_{i'}^{A\dagger}\otimes|i^B\rangle\langle i'^B|] \\ &= \sum_i K_i^A|\psi^A\rangle\langle\psi^A|K_i^{A\dagger} \\ &= \mathcal{E}^A(|\psi^A\rangle\langle\psi^A|)\end{aligned} \tag{16.26}$$

through the unitary transformation U^{AB} of the composite system, into which it is also transformed by the quantum operation $\mathcal{E}^A$. Every trace-conserving quantum operation on $\mathcal{H}^A$ therefore has a *unitary implementation* on $\mathcal{H}^A\otimes\mathcal{H}^B$ in the form

$$\mathcal{E}^A(\rho^A) = \mathrm{tr}_B[U^{AB}(\rho^A\otimes|0^B\rangle\langle 0^B|)U^{AB\dagger}]\,. \tag{16.27}$$

For every trace-conserving quantum operation $\mathcal{E}^A$ of a system S^A, a composite system S^{AB} with the Hilbert space $\mathcal{H}^A\otimes\mathcal{H}^B$, an initial state $|0^B\rangle$ of S^B, and a unitary transformation U^{AB} on $\mathcal{H}^{AB}$ can be found such that the subsystem S^A undergoes the evolution $\mathcal{E}^A$. U^{AB} is the extension of the operator $\hat{U}^{AB}$ defined in Eq. (16.24).

Quantum operations which do not conserve the trace In the case of incomplete quantum operations ($\sum_i K_i^{A\dagger}K_i^A < \mathbb{1}$), we complete the set of operators K_i^A by adding an operator K_* such that a complete set

$$\sum_i K_i^{A\dagger}K_i^A + K_*^{A\dagger}K_*^A = \mathbb{1} \tag{16.28}$$

is obtained. The space $\mathcal{H}^B$ is replaced by the space $\mathcal{H}^{B'}$ whose dimension is larger by one. An operator $U^{AB'}$ is constructed as above. Following the evolution with $U^{AB'}$, a projection is carried out using the projector P^B on $\mathcal{H}^B$. We thus replace U^{AB} in the above calculation by $P^BU^{AB'}$. Then in the equation which is analogous to Eq. (16.24), precisely the operator K_*^A vanishes. It also no longer occurs in the operator-sum representation (16.26).

16.3 Implementation of a Completely General Selective Measurement by Unitary Transformation and Projection*

We want to physically implement *most general selective measurements*. In these interventions, the state ρ^A is transformed on the occurrence of the measurement result m into the state $\tilde{\rho}_m'^A$, which is given by a quantum operation with superoperators $\mathcal{M}_m^A$

$$\rho^A \to \tilde{\rho}_m'^A = \mathcal{M}_m^A(\rho^A)\,. \tag{16.29}$$

*The sections marked with an asterisk * can be skipped over in a first reading.

The generalised measurement discussed in Sect. 13.3 is the special case $\mathcal{M}_m^A(\rho) = M_m^A \rho^A M_m^\dagger$.

The measurement outcome m is assumed to occur with a probability

$$p(m) = \mathrm{tr}_A[\tilde{\rho}_m'^A] = \mathrm{tr}_A[\mathcal{M}_m^A(\rho)] \,. \tag{16.30}$$

Thus for arbitrary ρ^A,

$$\mathrm{tr}_A[\sum_m \mathcal{M}_m^A(\rho^A)] = \sum_m p(m) = 1 \,. \tag{16.31}$$

The superoperators $\mathcal{M}_m^A$ have an operator-sum decomposition

$$\mathcal{M}_m^A(\rho) = \sum_i M_{m,i}^A \rho^A M_{m,i}^{A\dagger} \,. \tag{16.32}$$

The range of values of i can depend on m. As a result of Eq. (16.31), the ensemble must obey the completeness relation

$$\sum_{m,i} M_{m,i}^{A\dagger} M_{m,i}^A = \mathbb{1} \,. \tag{16.33}$$

We will make use of this formal analogy with Eq. (16.23).

A physical implementation can in principle also be found for this quite general measurement on the system A. To this end, we again introduce an ancilla system S^B. The dimension of the associated Hilbert space $\mathcal{H}^B$ is taken to be equal to the number of all the Kraus operators employed, $M_{m,i}^A$. In $\mathcal{H}^B$, we choose an ONB $\{|m, i^B\rangle\}$. Then we can follow the procedure of Sect. 16.2 step by step. Let the initial state of the composite system be $|\psi^A, 0^B\rangle$. An operator $\hat{U}^{AB}$ with the action

$$\hat{U}^{AB}|\psi^A, 0^B\rangle := \sum_{m,i} M_{m,i}^A |\psi^A\rangle |m, i^B\rangle \tag{16.34}$$

is defined. It can be extended by again using Eq. (16.33) to a unitary operator U^{AB} on the entire space $\mathcal{H}^A \otimes \mathcal{H}^B$.

We introduce orthogonal projection operators P_m^B on the ancilla system S^B belonging to the measurement outcome m:

$$P_m^B := \sum_i |m, i^B\rangle\langle m, i^B| \,, \quad \sum_m P_m^B = \mathbb{1}^B \,. \tag{16.35}$$

A unitary overall transformation with subsequent projection P_m^B on the ancilla system S^B yields

$$\begin{aligned} P_m^B U^{AB} |\psi^A\rangle |0^B\rangle &= \left(\sum_{i'} |m, i'^B\rangle\langle m, i'^B| \right) \left(\sum_{n,i} M_{n,i}^A |\psi^A\rangle |n, i^B\rangle \right) \\ &= \sum_i M_{m,i}^A |\psi^A\rangle |m, i^B\rangle \,. \end{aligned} \tag{16.36}$$

To determine the reduced density operator of the outgoing system S^A, we proceed as in Eq. (16.26) and obtain with Eq. (16.32) the desired relation (16.29)

$$\tilde{\rho}'^A_m = \sum_i M^A_{m,i} |\psi^A\rangle\langle\psi^A| M^{A\dagger}_{m,i} = \mathcal{M}^A_m(|\psi^A\rangle\langle\psi^A|) \,. \tag{16.37}$$

The probability $p(m)$ with which the measurement result m occurs in the projective measurement on S^B is given by the reduced density operator ρ^B of the ancilla system S^A after the unitary evolution. By taking the partial trace tr_A in Eq. (16.36), we first obtain ρ^B after an intermediate computation which we do not show here, and from it, $p(m)$:

$$p(m) = \mathrm{tr}_B[P^B_m \rho^B P^B_m] \,. \tag{16.38}$$

One can then read off the explicit expressions which we will not write out here, and with the aid of Eq. (16.37), obtain the final relation required (compare Eq. (16.30))

$$p(m) = \mathrm{tr}_A[\mathcal{M}^A_m(|\psi^A\rangle\langle\psi^A|)] \,. \tag{16.39}$$

Because of the linearity of the superoperators, the calculation can be applied to density operators.

We obtained the following result: *a completely general selective measurement on the system S^A, yielding the measurement outcome m and the change of state $\mathcal{M}_m(\rho)$, can be implemented as a unitary evolution U^{AB} of the composite system S^{AB} (extended to include S^B), with a subsequent projective measurement on S^B by the projector P^B_m. The probability that the measurement outcome m is found in the projective measurement on S^B agrees with the probability $p(m)$ that the measurement operation $\mathcal{M}^A_m$ acts on the state of S^A. U^{AB} and P^B_m are here given by Eqs. (16.34) and (16.35).* This represents a reduction to projective measurements, which cannot be physically further reduced. For details of the operational procedure of selection, we refer to Sect. 7.4.2.

16.4 Complementary Topics and Further Reading*

- On the proofs: [Kra 83], [Sch 96b]. See also the literature references in Sect. 14.5.
- Unitary evolutions are reversible. Under which conditions are quantum operations reversible? Compare [NCS 98].

16.5 Problems for Chapter 16

Prob. 16.1 [for 16.1]: Prove Eq. (16.8).

Prob. 16.2 [for 16.3]: Complete the intermediate calculation which leads to Eq. (16.38) and to Eq. (16.39).

*The sections marked with an asterisk * can be skipped over in a first reading.

© 2007 WILEY-VCH Verlag GmbH & Co.

References

Reference categories

- It all began with these articles: [EPR 35], [Sch 35].
- The classical reference: [vNe 68].
- A selection of general textbooks on quantum mechanics: [Pri 81], [Sak 85], [Ish 95], [Bal 98], [CDL 05].
- Books with more or less complete treatments of the field (in part including the experimental aspects): [Per 93], [d'Es 95], [Pre 98], [d'Es 99], [Aul 00], [NC 00], [Bru 03], [SS 04], [SS 05], [MM 05], [LeB 06], [Dio 06].
- Books in which some subfields are treated in detail: [Jor 69], [Kra 83], [BLM 91], [BGL 95], [Bub 97], [Hom 97], [Mit 98], [MW 98], [WC 98], [AS 99], [Bro 99], [Gru 99], [Pit 00], [CP 01], [Hir 01], [Hol 01], [Gri 02], [BCS 04], [SS 04].
- A collection of review articles: [GJK 96], [LSP 98], [Bra 99], [BEZ 00], [BGJ 00], [ABH 01], [BZ 02], [DM 02], [CB 02], [Hei 02], [Lom 02], [LB 02], [PR 04].
- Review articles in journals or books with complete treatments of the field: [Ben 95], [Joz 98], [PV 98], [Ste 98], [BD 00], [Lal 01], [Key 02], [KLB 04], [Bub 05b].
- Review articles in journals or books in which a subfield is treated in detail: [DiV 98], [Ste 98], [WW 01], [GM 02], [GRT 02], [Ved 02], [Zur 03], [PT 04], [Sch 04], [SIG 05].
- Review articles on experiments relevant to quantum information theory and collections which contain such articles: [LSP 98], [Bra 99], [BEZ 00] [ABH 01], [BZ 02], [DM 02], [GRT 02], [Hei 02], [VC 04].
- A collection of problems and solutions: [SH 04].
- Collections of elementarised review articles: [AM 96], [Aud 06].
- Treatments of the field for school pupils: [KM 02].
- Resource letters: [DG 71], [Bal 87].
- A very complete classified bibliography: [Cab 04].

Bibliography

[ABH 01] Alber, G., Beth, T., Horodecki, M., Horodecki, P., Horodecki, R., Rötteler, M., Weinfurter, H., Werner, R. und Zeilinger, A.: *Quantum Information – An Introduction to Basic Theoretical Concepts and Experiments*, Springer-Verlag, Berlin 2001.

[ADK 03] Audretsch, J., Diósi, L. und Konrad, T.: *Estimating the postmeasurement state*, Phys. Rev. **A68**, 034302 (2003).

[AS 99] Afriat, A. und Selleri, F.: *The Einstein, Podolsky and Rosen paradox in atomic, nuclear, and particle physics*, Plenum Press, New York, 1999.

[AKK 04] Audretsch, J., Klee, F. E. und Konrad, T.: *Monitoring Quantum Oscillations with very small Disturbance*; quant-ph/0408107 (2004).

[AKS 02] Audretsch, J., Konrad, T. und Scherer, A.: *Quantum optical weak measurements can visualize photon dynamics in real time*, Phys. Rev. **A65**, 033814(1-6) (2002).

[AL 91] Audretsch, J. und Lämmerzahl, C.: *Establishing the Riemannian structure of space-time by means of light rays and free matter waves*, J. Math. Phys. **32**, 2099-2105 (1991).

[AM 88] Audretsch, J. und Mainzer, K. (Ed.): *Philosophie und Physik der Raum-Zeit*, BI-Wiss.-Verl., Mannheim, 1988.

[AM 94a] Audretsch, J. und Müller, R.: *Localized discussion of stimulated processes for Rindler observers and accelerated detectors*, Phys. Rev. **D49**, 4056-4065 (1994).

[AM 94b] Audretsch, J. und Müller, R.: *Radiation from a uniformly accelerated particle detector: Energy, particles and the quantum measurement process*, Phys. Rev. **D49**, 6566-6575 (1994).

[AM 96] Audretsch, J. und Mainzer, K.(Ed.): *Wieviele Leben hat Schrödingers Katze? – Zur Physik und Philosophie der Quantenmechanik*, Spektrum Akad. Verlag, Heidelberg, 1996.

[ANZ 02] Arndt, M., Nairz, O. und Zeilinger, A.: *Interferometry with Macromolecules and the Mesoscopic World*, in [BZ 02], p. 334–350, 2002.

[Asp 99] Aspect, A.:*Bell's inequality test: More ideal than ever*, Nature **398**, 189-190 (1999).

[Asp 02] Aspect, A.: *Bell's Theorem: The Naive View of an Experimentalist*, in [BZ 02], 119–153 (2002).

[Aud 06] Audretsch, J. (Ed.): *Entangled World – The Fascination of Quantum Information and Computation*, Wiley-VCH, Berlin, 2006. Erweiterte Fassung der deutschen Ausgabe: *Verschränkte Welt – Faszination der Quanten*, Wiley-VCH, Weinheim, 2002.

[Aud 06a] Audretsch, J.: *View into the quantum world I: fundamental phenomena and concepts*, in [Aud 06], p. 1-37, *View into the quantum world II: entanglement and its consequences*, in [Aud 06], p. 39-62, 2006.

[Aul 00] Auletta, G.: *Foundations and Interpretations of quantum mechanics*, World Scientific, Singapore, 2000.

[Bal 70] Ballentine, L. E.: *The Statistical Interpretation of Quantum Mechanics*, Rev. Mod. Phys. **42**, 358–381, (1970).

[Bal 87] Ballentine, L.E.: *Resource letter IQM-2: Foundations of quantum mechanics since the Bell inequalities*, Am. J. Phys. **55**, 785-792 (1987).

[Bal 98] Ballentine, L. E.: *Quantum mechanics – a modern development*, World Scientific, Singapore, 1998.

[Bar 95] Barenco, A.: *A universal two-bit gate for quantum computation*, Proc. R. Soc. Lond. **A449**, 679-683 (1995).

[Bar 96] Barenco, A.: *Quantum physics and computers*, Contemp. Phys. **37**, 375–389 (1996).

[Bar 98] Barenco, A.: *Quantum computation: an introduction*, in [LSP 98], p. 143–183 (1998).

[Bar 00] Barrett, J.: *Everett's Relative-State formulation of Quantum Mechanics*, Stanford Encyclopedia of Philosophy, http://plato.stanford.edu/entries/qm-everett, 2000.

[BB 84] Bennett, C.H. und Brassard, G.: *Quantum key distribution and coin tossing*, in *Proc. of IEEE Int. Conf. on Computers, Systems and Signal Processing (Bangalore, India, 1984)*, p. 175-179, IEEE, New York, 1984.

[BBC 93] Bennett, C. H., Brassard, G., Crépeau, C., Jozsa, R., Peres, A. und Wooters, W.K.: *Teleporting an unknown quantum state via dual classical and EPR channels*, Phys. Rev. Lett. **70**, 1895-1899 (1993).

[BBC 95] Barenco, A., Bennet, C. H., Cleve, R., DiVincenzo, D. P., Margolus, N., Shor, P. W., Sleator, T., Smolin, J. A. und Weinfurter, H.: *Elementary gates for quantum computation*, Phys, Rev. **A52**, 3457-3467 (1995); quant-ph/9503016.

[BBD 97] Boschi, B., Branca, S., De Martini, F. und Hardy, L.: *A Ladder Proof of Nonlocality Without Inequalities: Theoretical and Experimental Results*, Phys. Rev. Lett. **79**, 2755-2758 (1997).

[BBH 98] Boyer, M., Brassard, G., Høyer, P. und Tapp, A.: *Tight bounds on quantum searching*, Fortschr. Phys. **46**, 493–505 (1998); quant-ph/9605034.

[BBM 92] Bennett, C. H., Brassard, G. und Mermin, N. D.: *Quantum Cryptography without Bell's Theorem*, Phys. Rev. Lett. **68**, 557–559 (1992).

[BBP 96] Bennett, Ch. H., Brassard, G., Popescu, S., Schumacher, B., Smolin, J. A. und Wootters, W. K.: *Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels*, Phys. Rev. Lett. **76**, 722-725 (1996).

[BCS 04] Berenti, G., Casati, G. und Strini, G.: *Principles of Quantum Computation, Vol. I: Basic Concepts*, World Scientific, Singapore, 2004.

[BeB 96] Bennett, Ch. H ., Bernstein, H. J., Popescu und S., Schumacher, B.: *Concentrating partial entanglement by local operations*, Phys. Review **A 53**, 2046–2052 (1996).

[BC 81] Beltrametti, E. G. und Cassinelli, G.: *The Logic of Quantum Mechanics*, Addison-Wesley, Reading 1981.

[BCZ 99] Briegel, H.-J., Cirac, I. und Zoller, P.: *Quantencomputer: wie Verschränkung in der Informationsverarbeitung verwendet werden kann*, Phys. Blätter **55**, Nr. 9, 37–43 (1999).

[BD 00] Bennett, Ch. H. und DiVincenzo, D. P.: *Quantum information and computation*, Nature **404**, 247–255 (2000).

[BDF 99] Bennett, C. H., DiVincenzo, D. P., Fuchs, C. A., Mor, T., Rains, E.M., Shor, P. W., Smolin, J. A. und Wooters, W. K.: *Quantum nonlocality without entanglement*, Phys. Rev. **A59**, 1070-1091 (1999); quant-ph/9804053. *Quantum Repeaters: The Role of Imperfect Local Operations in Quantum Communication*, Physical Review Letters **81**, 5932–5935 (1998); /quant-ph/9803056.

[BDM 98] Berman, G. P., Doolen, G. D., Mainieri, R. und Tsifrinovich, V. I.: *Introduction to quantum computers*, World Scientific, Singapore, 1998.

[Bel 64] J. S. Bell: *On the Einstein-Podolsky-Rosen paradox*, Physics **1**, 195–200 (1964). Nachdruck in J. S. Bell *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, Cambridge, 1987.

[Ben 92] Bennett, C. H.: *Quantum cryptography using any two nonorthogonal states*, Phys. Rev. Lett., **68**, 3121–3124 (1992).

[Ben 95] Bennett, C. H.: *Quantum Information and Computation*, Physics Today, October 24–30 (1995).

[Ber 97] Berman, P. R.: *Atom Interferometry*, Academic Press, San Diego, 1997.

[BEZ 00] Bouwmeester, D., Ekert, A. und Zeilinger, A. (Ed.): *The physics of quantum information – quantum cryptography, quantum teleportation, quantum computation*, Springer-Verlag, Berlin, 2000.

[BGJ 00] Blanchard, Ph., Giulini, D., Joos, E., Kiefer, C. und Stamatescu, J.-O. (Ed.): *Decoherence: Theoretical, Experimental, and Conceptual Problems*, Springer, Berlin, 2000.

[BGL 95] Busch, P., Grabowski, M. und Lahti, P. J.: *Operational quantum mechanics*, Springer-Verlag, Berlin, 1995. Phys. Rev. **A 54**, 1844–1852 (1996)

[BHL 02] Bouwmeester, D., Howell, J. C. und Lamas-Linares, A.: *Quantum Information Science Using Photons*, in [Hei 02], p. 149–197 (2002).

[Bla 06] Blatt, R.: *Quantum information processing: Dream and realization*, in [Aud 06], p. 235 - 270, 2006.

[BLM 91] Busch, P., Lahti, P. J. und Mittelstaedt, P.: *The Quantum Theory of Measurement*, Springer-Verlag, Berlin, 1991.

[Boh 51] Bohm, D.: *Quantum Physics* Prentice Hall, New York, 1951.

[Boh 85] Bohr, N.: *Collected Works*, Vol. 6, p.296, North Holland, Amsterdam, 1985.

[Boh 83] Bohm, D. J.: *A Suggested Interpretation of the Quantum Theory in Terms of "Hidden" Variables, I and II* in Wheeler, J. A. und Zurek, W. H. (eds.): *Quantum theory and measurement*, Princeton University Press, Princeton, New Jersey, 1983.

[BP 02] Breuer, H.-P. und Petruccione, F.: *The Theory of Open Quantum Systems*, Oxford University Press, Oxford, 2002. *Multi-Partial Entanglement*, in [BEZ 00] S. 197–209, (2000).

[Bra 99] Braunstein, P. (Ed.): *Quantum computing: Where do we want to go tomorrow*, Wiley-VCH, Weinheim, 1999.

[Bra 99a] Braunstein, S.: *Quantum Computation*, in [Bra 99], p. 1–21, 1999.

[Bra 02] Brandt, H. E.: *Qubit Devices* in [Lom 02], S. 67–139 (2002).

[Bro 99] Brooks, M. (Ed.): *Quantum computing and communications*, Springer-Verlag, Berlin, 1999.

[Bru 02] Bruß, D.: *Characterizing Entanglement*, J. Math. Phys. **43**, 4237-4251 (2002), quant-ph/0110078 (2001).

[Bru 03] Bruß, D.: *Quanteninformation*, S. Fischer Verlag, Frankfurt a.M., 2003.

[Bub 97] Bub, J.: *Interpreting the Quantum World*, Cambridge University Press, Cambridge, 1997.

[Bub 05a] Bub, J.: *Quantum Theory is about Quantum Information*, Foundations of Physics, **35**, 541-560 (2005).

[Bub 05b] Bub, J.: *Quantum Information and Communication*, quant-ph/0512125v1.

[Bus 97] Busch, P.: *Is the quantum state (an) observable?* in: R.S. Cohen et. Al. (Ed.) *Potentiality and Passion -at-a- Distance*, S. 61–70, Kluwer Acad. Publ., 1997; quant-ph/9604014.

[Bus 99] Busch, P.: *Quantum States as Effect Valuations: Giving Operational Content to von Neumann's No-Hidden-Variable Theorem*; quant-ph/9909073

[Bus 02] Busch, P.: *Classical versus quantum ontology*, Studies in History and Philosophy of Modern Physics **33**, 517–539 (2002).

[BVK 98] Bose, S., Vedral, V. und Knight, P. L.: *Multiparticle generalization of entanglement swapping*, Phys. Review **A 57**, 2, 822–829 (1998); quant-ph/9708004.

[BvL 05] Braunstein, S.L. und van Loock, P.: *Quantum Information with continuos variables*, Rev. Mod. Phys. **77**, 513-577 (2005).

[BVS 96] Bennett, CH. H., DiVincenzo, D. P., Smolin, J. A. und Wootters, W. K.: *Mixed-state entanglement and quantum error correction*, Phys. Rev. **A 54**, 3824-3851 (1996).

[BW 92] Bennett, C. und Wiesner, S.: *Communication via one- and two particles operators on Einstein-Podolsky-Rosen states*, Phys. Rev. Lett. **69**, 2881-2884 (1992).

[BZ 02] Bertelmann, R. A. und Zeilinger, A. (Ed.): *Quantum [un]speakables – From Bell to quantum information*, Springer-Verlag, Berlin, 2002.

[Cab 04] Cabello, A.: *Bibliographic guide to the foundations of quantum mechanics and quantum information*; quant-ph/0012089v12 (2004).

[Car 99] Carmichael, H.J.: *Statistical Methods in Quantum Optics 1 – Master Equations and Fokker-Planck Equations*, Springer, Berlin, 1999.

[CB 02] Chen, G. und Brylinski, R.K. (Ed.): *Mathematics of Quantum Computation*, Chapman and Hall/CRC, Boca Raton, 2002.

[CBH 03] Clifton, R., Bub, J. und Halvorson, H.: *Characterizing Quantum Theory in Terms of Information Theoretic Constraints*, Found. Phys. **33**, 1561-1591 (2003).

[CD 94] Caves, C. M. und Drummond, P.D.: *Quantum limits on bosonic communication rates*, Rev. Mod. Phys. **66**, 481-537 (1994).

[CDL 05] Cohen-Tannoudji, C., Diu, B. und Laloë, F.: *Quantum Mechanics, Vol. 1*, Wiley, New York, 2005.

[CEM 98] Cleve, R., Ekert, A., Macchiavello, C. und Mosca, M.: *Quantum algorithms revisited*, Proc. R. Soc. London, **A454**, 339–354 (1998).

[CF 96] Caves, C. M. und Fuchs, Ch. A.: *Quantum Information: How much Information in a State Vektor?*; quant-ph/9601025

[CHS 69] Clauser, J. F. , Horne, M. A., Shimony, A. und Holt, R. A.: *Proposed experiment to test local hidden-variable theories*, Phys. Rev. Lett., **23**, 880–884 (1969).

[Cir 02] Cirac, J. I.: *Quantum Information: Entanglement, Purification, Error Correction and Quantum Optical Implication*, in [Hei 02], p. 199–239, 2002.

[CLK 00] Cory, D. G., Laflamme, R., Knill, E., Viola, L., Havel, T. F., Boulant, N., Boutis, G., Fortunato, E., Lloyd, S., Martinez, R., Negrevergne, C., Pravia, M., Sharf, Y., Teklemariam, G., Weinstein, Y. S. und Zurek, W. H.: *NMR based quantum information processing: Achievements and prospects*, Fortschr. Phys. **48**, 9–11 (Special issue: Experimental proposals for quantum computation), 875–907 (2000); quant-ph/0004104.

[CM 91] Carnal, O. und Mlynek, J.: *Young's double-slit experiment with atoms: A simple atom interferometer*, Phys. Rev. Lett. **66**, 2689-2692 (1991).

[CP 01] Calude, C. S. und Paun, G.: *Computing with Cells and Atoms – An Introduction to quantum, DNA and membrane computing*, Taylor and Francis, London, 2001.

[CST 89] Campos, R. A., Saleh, B. E. A. und Teich, M. C.: *Quantum-mechanical lossless beam splitter: SU(2) symmetry and photos statistics*, Phys. Rev. **A 40**, 1371–1384 (1989).

[CU 99] Crell, B. und Uhlmann, A.: *Einführung in die Grundlagen und Protokolle der Quanteninformatik*, *www.uni-leipzig.de/ ntz/abs/papers/ntz3398.ps*

[d'Es 95] D'Espagnat, B.: *Veiled reality – an analysis of present-day quantum mechanical concepts*, Addison-Wesley, Reading (Mass.), 1995.

[d'Es 99] D'Espagnat, B.: *Conceptual foundations of quantum mechanics*, Perseus Books, Reading, 1999.

[Dav 76] Davies, E. B.: *Quantum Theory of open systems*, Academic Press, New York, 1976.

[DE 98] Deutsch, D. und Ekert, A. K.: *Quantum computation*, Phys. World **11**, 53–57 (1998).

[DEJ 96] Deutsch, D., Ekert, A., Josza, R., Macchiavello, C., Popescu, S. und Sanpera, A.: *Quantum privacy amplification and the security of quantum cryptography*, Phys. Rev. Lett. **77**, 2818-2821 (1996).

[Deu 85] : Deutsch D.: *Quantum theory, the Church-Turing principle and the universal quantum computer* , Proc. R. Soc. Lond. **A 400**, 97–117 (1985).

[Deu 89] Deutsch, D.: *Quantum computational networks*, Proc. R. Soc. Lond., **A 425**, 73–90 (1989).

[Deu 96] Deutsch, D.: *Comment on the many minds interpretation of quantum mechanics*, Brit. J. Philos. Sci. **47**, 222 (1996).

[DG 71] DeWitt, B. S. und Graham, R. N.: *Resource Letter IQM-1 on the Interpretation of Quantum Mechanics*, Am. J. Phys. **39**, 724–738 (1971).

[DG 73] DeWitt, B. S. und Graham, R. N. (Ed.): *The many-worlds interpretation of quantum mechanics*, Princeton University Press, Princeton, New Jersey, 1973.

[DG 98] Duan, L.M. und Guo, G.-C.: *Probabilistic cloning and identification of linearly independent quantum states*, Phys. Rev. Lett. **80**, 4999-5002 (1998); quant-ph/9804064.

[DHR 02] Donald, M. J., Horodecki, M. und Rudolph, O.: *The uniqueness theorem for entanglement measures*; quant-ph/0105017 v2 (2002).

[Die 82] Dieks, D.: *Communication by EPR Devices*, Phys. Lett. **A 92**, 271–272 (1982).

[Dio 06] Diósi, L.: *Short Course on Quantum Information Theory, an approach from theoretical physics*, Springer, Berlin, to appear.

[DiV 95] DiVincenzo, D. P.: *Two-bit gates are universal for quantum computation*, Phys. Rev. A **51**, 1015-1022 (1995); cond-mat/9407022.

[DiV 98] DiVincenzo, D. P.: *Quantum gates and circuits*, Proc. R. Soc. London, **A 454**, 261–276 (1998).

[DiV 00] DiVincenzo, D. P.: *The physical implementation of quantum computation*, Fortschr. Phys. **48**, 9–11 (Special issue: Experimental proposals for quantum computation), 771–783 (2000); quant-ph/0002077.

[DJ 92] Deutsch D., Jozsa R.: *Rapid solution of problems by quantum computation* , Proc. R. Soc. Lond. **A 439**, 553–558 (1992).

[DM 02] DeMartini, F. und Monroe, C.: *Experimental Quantum Computation and Information*, „Enrico Fermi“ Course CXLVIII, IOS Press, Amsterdam, 2002.

[DMB 97] Di Giuseppe, G., De Martini, F. und Boschi, D.: *Experimental Test of the Violations of Local Realism in Quantum Mechanics without Bell Inequalities*, Phys. Rev. **A56**, 1-6 (1997).

[DNR 98] Dürr, S., Nonn, T. und Rempe, G.: *Fringe visibility and which-way information in an atom interferometer*, Phys. Rev. Lett. **81**, 5705–5709 (1998).

[DR 00] Dürr, S. und Rempe, G.: *Wave-Particle Duality in an Atom Interferometer*, Advances Atomic, Molecular and Optical Physics, **45**, 29–71 (2000).

[EB 69] Einstein, A., Born, H. und M.: *Briefwechsel 1916-1955*, Nymphenburgische Verlagsbuchhandlung, München 1969. Reprinted in *The Born-Einstein Letters*, Walker, New York, 1971.

[EGH 00] Ekert, A., Gisin, N., Huttner, B., Inamori, H. und Weinfurter, H.: *Quantum Cryptography*, in [BEZ 00], p. 15–48 (2000).

[EHI 01] Ekert, A., Hayden, P. M. und Inamori, H.: *Basic concepts in quantum computation*, in Les Houches, Session LXXII *Coherent atomic matter waves*, Kaiser, R., Westbrook, C. and David, F. (Ed.), p. 661–701, Springer, Berlin, 2001.

[EJ 96] Ekert, A. und Jozsa, R.: *Quantum computation and Shor's factoring algorithm*, Rev. Mod. Phys. **68**, 733–753 (1996).

[Eke 91] Ekert, A. K.: *Quantum cryptography based on Bell's theorem*, Phys. Rev. Lett. **67**, 661-663 (1991).

[EMY 02] Eckert,K., Mompart, J., Yi, X. X., Schliemann, J., Bruß, D., Birkl, G. und Lewenstein, M.: *Quantum computing in optical microtraps based on the motional states of neutral atoms*, Phys. Rev. **A66**, 042317 (2002).

[Eng 99] Engler, B.-G.: *Remarks on some basic issues in quantum mechanics*, Zeitschrift f. Naturforschung **54**, 11–32 (1999).

[EPR 35] Einstein, A., Podolsky, B. und Rosen, N.: *Can quantum-mechanical description of reality be considered complete?*, Phys. Rev. **47**, 777–780 (1935).

[ESB 02] Eckert, K., Schliemann, J., Bruß, D. und Lewenstein, M.: *Quantum Correlations in Systems of Indistinguishable Particles*, Ann. Phys. **299**, 88-127 (2002); quant-ph/020360

[Esf 02] Esfeld, M.: *Einführung in die Naturphilosophie*, Wissenschaftliche Buchgesellschaft, Darmstadt, 2002.

[Esf 04] Esfeld, M.: *Quantum entanglement and a metaphysics of relations*, Stud. Hist. Phil. of Science Part B, Vol. **35**, Issue 4, 625–641 (2004).

[Esf 06] Esfeld, M.: *Quantum theory: a challenge for philosophy*, in [Aud 06] p. 271–296, 2006.

[ESW 99] Englert, B.-G., Scully, M. O. und Walther, H.: *Quantum erasure in double-slit interferometers with which-way detectors* Am. J. Phys. **67**, 325–329 (1999).

[EV 93] Elizur, A.C. und Vaidman, L.: *Quantum Mechanical Interaction-Free Measurements*, Found. Phys. **23**, 987-997 (1993).

[Eve 57] Everett, H.: *Relative state formulation of quantum mechanics*, Rev. Mod. Phys. **29**, 454–462 (1957).

[Fan 57] Fano, U.: *Description of States in Quantum Mechanics by Density Matrix and Operator Techniques*, Rev. Mod. Phys. **29**, 74–93 (1957)

[FK 98] Fischer, H. und Kaul, H.: *Mathematik für Physiker Bd. 2: Gewöhnliche und partielle Differentialgleichungen, mathematische Grundlagen der Quantenmechanik*, Teubner, Stuttgart, 1998.

[Fle 00] Fleming, G. N.: *Operational Quantum Physics*, Stud. Hist. Phil. Mod. Phys. **31**, 117–125 (2000).

[FP 00] Fuchs, C. A. und Peres, A.: *Quantum theory needs no interpretation*, Physics Today, March 2000, 70–71 sowie Sept. 2000, 11–14, 1990.

[FPC 00] Friedman, J. R., Patel, V., Chen, W., Tolpygo, S. K. und Lukens, J. E.: *Quantum superposition of distinct macroscopic states*, Nature **406**, 43–46 (2000).

[Fuc 03] Fuchs, Ch.: *Quantum Mechanics as Quantum Information, mostly*, Journ. mod. Optics **50**, 987-1023 (2003); quant-ph/0205039 (2002).

[GHZ 89] Greenberger, D. M., Horne, M. A. und Zeilinger, A.: *Going Beyond Bell's Theorem* in M. Kafatos (Ed.) *Bell Theorem, Quantum Theory and Conceptions of the Universe*, S. 69–72, Kluwer, Dosdrecht, 1989.

[GHZ 90] Greenberger, D. M., Horne, M. A, Shimony, A. und Zeilinger, A.: *Bell's Theorem without inequalities*, Am. J. Phys. **58**, 1131–1143 (1990).

[GJK 96] Giulini, D., Joos, E., Kiefer, C., Kupsch, J., Stamatescu, I.-O. und Zeh, D. H.: *Decoherence and the appearance of the classical world in quantum theory*, Springer-Verlag, Berlin, 1996.

[Gle 57] Gleason, A. M.: *Measures on the Closed Subspaces of a Hilbert Space*, J. Math. Mech. **6**, 885–894 (1957)

[GM 02] Galindo, A. und Martín-Delgado, M. A.: *Information and computation: Classical and quantum aspects*, Rev. Mod. Phys. **74**, 347–423 (2002), quant-ph/0112105.

[Gri 02] Griffiths, R. B.: *Consistent Quantum Theory*, Cambridge University Press, Cambridge UK, 2002.

[Gro 96] Grover, L. K.: *A fast quantum mechanical algorithm for database search* in *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing*, p. 212–219, Philadelphia, 1996.

[GRT 02] Gisin, N., Rikordy, G., Tittel, W. und Zbinder, H.: *Quantum cryptography*, Rev. Mod. Phys. **74**, 145 – 195 (2002).

[Gru 99] Gruska, J.: *Quantum Computing*, McGraw-Hill, London, 1999.

[GRW 86] Ghirardi, G. C., Rimini, A. und Weber, T.: *Unified dynamics for microscopic and macroscopic systems*, Phys. Rev. **D34**, 470-491 (1986).

[HAD 95] Hughes, R. J., Alde, D. M., Dyer, P., Luther, G. G., Morgan, G. L. und Schauer, M.: *Quantum Cryptography*, Contemp. Phys. **36**, 149–163 (1995); quant-ph/9504002.

[Har 01a] Hardy, L.: *Why Quantum Theory?*; quant-ph/011068v1 (2001).

[Har 01b] Hardy, L.: *Quantum Theory from five reasonable Axioms*; quant-ph/0101012v4 (2001).

[Har 92] Hardy, L.: *Quantum Dynamics, Realistic Theories and Lorentz-Invariant Realistic Theories*, Phys. Rev. Lett. **68**, 2981-2984 (1992).

[Har 93] Hardy, L.: *Nonlocality for Two Particles without Inequalities for Almost All Entangled States*, Phys. Rev. Lett. **71**, 1665-1668 (1993).

[Har 98] Hardy, L.: *Spooky action at a distance in quantum mechanics* Contemp. Physics, **39**, 419-429 (1998).

[HE 02] Horodecki, P. und Ekert, A.K.: *Method for direct detection of quantum entanglement*, Phys. Rev. Lett. **89**, 127902 (2002).

[Hea 99] Healey, R.: *Holism and Nonseperability in Physics*, in *Stanford Encyclopedia of Physics*, *http://plato.stanford.edu/entries/physics-holism*.

[Hei 02] Heiss, D. (Ed.): *Fundamentals of Quantum Information*, Springer-Verlag, Berlin, 2002.

[Hel 06] Held, C.: *The Bohr-Einstein debate and the fundamental problem of quantum mechanics*, in [Aud 06], p. 65-90, 2006.

[HHH 96] Horodecki, M., Horodecki, P. und Horodecki, R.: *Separability of mixed states: Necessary and sufficient conditions*, Phys. Lett. A **223**, 1-8 (1996); quant-ph/9605038.

[HHH 01] Horodecki, M., Horodecki, P. und Horodecki, R.: *Mixed State entanglement and Quantum Communication*, in [ABH 01], p. 151–195, 2001.

[Hir 01] Hirsenvalo, M.: *Quantum computing*, Springer-Verlag, Berlin, 2001.

[HK 69] Hellwig, H.-E. und Kraus, K.: *Pure operations and measurements*, Comm. math. Phys. **11**, 214–220 (1969).

[HK 70] Hellwig, H.-E. und Kraus, K.: *Operations and measurements*, Comm. math. Phys. **16**, 142–147 (1970).

[HN 99] Hughes, R. und Nordholdt, J.: *Quantum cryptography takes to the air*, Physics World, May, 31–35 (1999).

[Hol 01] Holevo, A. S.: *Statistical Structure of Quantum Theory*, Springer-Verlag, Berlin, 2001.

[Hom 97] Home, D.: *Conceptual foundations of quantum physics – an overview from modern perspective*, Plenum Press, New York, 1977.

[HRB 02] Haroche, S., Raimond, J. M. und Brune, M.: *Entanglement, complementority and decoherence in cavity QED experiments*, in [DM 02], p. 3–36, 2002.

[HS 91] Home, D. und Selleri, F.: *Bell's Theorem and the EPR Paradox*, Riv. Nuov. Cim. **14**, n9. 1–95 (1991).

[HUH 03] Hackermüller, L., Uttenthaler, S., Hornberger, K., Reiger, E., Brezger, B., Zeilinger, A. und Arndt, M.: *Wave Nature of Biomolecules and Flourofullerenes*, Phys. Rev. Lett. **91**, 090408-1 – 09048-4 (2003).

[HW 79] Hardy, G. und Wright, E.: *An Introduction to the Theory of Numbers*, Clarendon Press, Oxford, 1979.

[HW 97] Hill, S. und Wooters, W. K.:*Entanglement of a pair of quantum bits*, Phys. Rev. Lett. **78**, 5022-5025 (1997).

[HWZ 87] Hellmuth, T., Walther, H., Zajonc, A. und Schleich, W.: *Delayed-choice experiments in quantum interference*, Phys. Rev. **A 35**, 2532-2541 (1987).

[HKW 95] Herzog, T. J., Kwiat, P. G., Weinfurter, H. und Zeilinger, A.:*Complementarity and the Quantum Eraser*, Phys. Rev. Lett. **75**, 3034-3037 (1995).

[IHB 90] Itano, W. H., Heinzen, D. J., Bollinger, J. J., und Wineland, D. J.: *Quantum Zeno Effect*, Phys. Rev. **A41**, 2295–2300 (1990).

[Ish 95] Isham, C.: *Lectures on quantum theory – mathematical and structural foundations*, Imperial College Press, distributed by World Scientific, Singapore, 1995.

[Joo 96] Joos, E.: *Decoherence Through Interaction with the Enviroment*, in [GJK 96], p. 35–136, 1996.

[Joo 06] Joos, E.: *Decoherence and the transition from quantum physics to classical physics*, in [Aud 06], p. 203-233, 2006.

[Jor 69] Jordan, Th. F.: *Linear Operators for Quantum Mechanics*, Wiley, New York, 1969.

[Joz 98] Jozsa, R.: *Quantum Information and its Properties*, in [LSP 98], p. 49–75 (1998).

[Joz 00] Jozsa, R.: *Quantum Algorithms*, in *The Physics of Quantum Information*, in [BEZ 00], p. 104–126, 2000.

[Joz 05] Jozsa, R.: *An introduction to measurement based quantum computation*; quant-ph/0508124.

[JS 94] Jozsa, R. und Schumacher, B.: *A new proof of the quantum noiseless coding theorem*, J. of Mod. Optics **41**, 2343–2349 (1994).

[Key 02] Keyl, M.: *Fundamentals of quantum information theory*, Physics Reports **369**, 431–548 (2002).

[KLB 04] Knill, E., Laflamme, R., Barnum, D., Dalvit, D., Dziarmaga, J., Gubernatis, L., Gurvits, L., Ortiz, G., Viola, L. und Zurek, W. H.: *Introduction to quantum Information Processing*; quant-ph/0207171 v1, 2002.

[KM 02] Küblbeck, J. K. und Müller, R.: *Die Wesenszüge der Quantenphysik – Modelle, Bilder, Experimente*, Aulis Verlag Deubner, Köln, 2002.

[Kon 03] Konrad, Th.: *On the theory and application of weak and unsharp measurements in Quantum Mechanics*, Doctoral thesis, University of Konstanz 2003, *http://www.ub.uni-konstanz.de/kops/volltexte/2003/1050/*.

[Kra 83] Kraus, K.: *States, Effects and Operations*, Springer-Verlag, Berlin, 1983.

[KSN 06] Konrad, Th., Scherer, A., Nock, M. und Audretsch, J.: *Heralded single-photon generation using imperfect single-photon sources and a two-photon absorbing medium*, Phys. Rev. **A73**, 032327 (2006); quant-ph/0602225.

[KSV 02] Kitaev, A. Y., Shen, A. H. und Vyalyi, M. N.: *Classical and Quantum Computation*, Amer. Math. Society, Providence, 2002.

[KWM 99] Kwiat, P. G., White, A. G., Mitchell, J. R., Nairz, O., Weihs, G., Weinfurter, H. und Zeilinger, A.:*High-Efficiency Quantum Interrogation Measurement via the Quantum Zeno Effect*, Phys. Rev. Lett. **83**, 4725-4728 (1999).

[Lal 01] Laloë, F.: *Do we really understand quantum mechanics? Strange correlations, paradoxes and theorems*, Am. J. Phys. **69**, 655–701 (2001).

[LB 02] Lomonaco, S. J. und Brandt, H. E. (Ed.): *Quantum computation and information*, American Mathematical Society, Providence, Rhode Island, 2002.

[LB 03] Leuchs, G. und Beth, T. (Ed.): *Quantum Information Processing*, Wiley-VCH, Weinheim, 2003.

[LB 39] London, F. und Bauer, E.: *La théorie de l'observation en mécanique quantique*, Hermann, Paris (1939). English translation: *The theory of observation in quantum mechanics*, in *Quantum Theory and Measurement*, pp. 217-259, ed. by Wheeler, J. A., Zurek, W. H., Princeton University Press, Princeton, 1983.

[LBC 00] Lewenstein, M., Bruß, D., Cirac, J. I., Kraus, B., Kuś, M., Samsonowicz, J., Sanpera, A. und Tarrach, R.: *Separability and distillability in composite quantum systems – a primer*, J. Mod. Opt. **77**, 2481-2499 (2000); quant-ph/0006064.

[LCS 99] Lütkenhaus, N., Calsamiglia, J. und Suominen, K.: *Bell measurements for teleportation*, Phys. Rev. **A59**, 3295-3300 (1999).

[LD 98] Loss, D. und DiVincenzo, D. P.: *Quantum computation with quantum dots*, Phys. Rev. A **57**, 120-126 (1998); cond-mat/9701055.

[LeB 06] Le Bellac, M.: *A Short Introduction to Quantum Information and Quantum Computing*, Cambridge Univ. Press, Cambridge, 2006.

[Leg 02] Leggett, A. J.: *Qubits, Cbits, Decoherence, Quantum Measurement and Enviroment*, in [Hei 02], p. 3–45, 2002.

[Leg 02a] Legget, A. J.: *Testing the limits of quantum mechanics: motivation, state of play, prospects*, J. Phys.: Condens. Matter **14**, R415-R451 (2002).

[LMP 98] Linden, N., Massar, S. und Popescu, S.: *Purifying noisy entanglement requires collective measurements*, Phys. Rev. Lett. **81**, 3279-3282 (1998); quant-ph/9805001.

[Lom 02] Lomonaco, S. J., (Ed.): *Quantum computation: A Grand mathematical challenge for the twenty-first century and the millenium*, American Mathematical Society, Providence, Rhode Island, 2002.

[Lom 02a] Lomonaco, S. J.: *Rosetta stone for Quantum Mechanics with an Introduction to Quantum Computation Version 1.5*, in [Lom 02], S. 3–65, 2002; quant-ph/0007045v1.

[LSP 98] Lo, H.-K., Spiller, T. und Popescu, S. (Ed.): *Introduction to Quantum Computing and Information*, World Scientific, Singapore 1998.

[Lud 55] Ludwig, G.: *Zur Deutung der Beobachtung in der Quantenmechanik*, Phys. Bl. 11, 489–494 (1955).

[Lud 83] Ludwig, G.: *Foundations of Quantum Mechanics I*, Springer-Verlag, New York, 1983.

[Lud 85] Ludwig, G.: *An Axiomatic Basis for Quantum Mechanics, Vol 1*, Springer, Berlin, 1985.

[Lud 89] Ludwig, G.: *Atoms: Are they real or are they objects*, Found. of Physics **19**, 971–983 (1989).

[Lud 90] Ludwig, G.: *Concepts of States in Physics*, Found. of Phys. **20**, 621–633 (1990).

[Lud 96] Ludwig, G.: *Die Katze ist tot* in [AM 96], p. 183–208, 1996.

[LW 03] Lidar, D. A. und Whaley, K. B.: *Decoherence-Free Subspaces and Subsystems* in *Irreversible Quantum Dynamics*, Benatti, F. und Floreanini (Ed.), p. 83–120, Springer, Berlin, 2003; quant-ph/0301032.

[Mai 96] Mainzer, K.: *Naturphilosophie und Quantenmechanik* in [AM 96], S. 245–299, 1996.

[Mar 78] Marlow, A. R. (Ed.): *Mathematical Foundation of Quantum Theory*, Academic Press, NY, 1978.

[MCK 05] Mintert, F., Carvalho, A.R.R., Kuas, M. und Buchleitner, A.: *Measures and dynamics of entangled states*; quant-ph/0505162 (2005).

[Mer 65] Merton, R. K.: *On the shoulders of Giants*, Free Press, New York, 1965.

[Mer 02] Mertens, S.: *Computational Complexity for Physicists*, Computing in Science and Engineering 4, 31–47 (2002); cond-mat/0012185v2.

[Mit 78] Mittelstaedt, P.: *Quantum logic*, Reidel, Dordrecht, 1978.

[Mit 96] Mittelstaedt, P.: *Objektivität und Realität in der Quantenphysik*, in [AM 96]. p. 125–155, 1996.

[Mit 98] Mittelstaedt, P.: *The interpretation of quantum mechanics and the measurement process*, Cambridge University Press, Cambridge, 1998.

[MM 05] Marinescu, D. C. und Marinescu, G. M.: *Approaching Quantum Computing*, Pearson Prentice Hall, Upper Saddle River, 2005.

[Mut 98] Mutschler, H.-D.: *Die Welt als Konstruktion* in Magerl, G. und Komarek, K. (Ed.) *Virtualität und Realität: Bild und Wirklichkeit der Naturwissenschaften*, Wien, 1998.

[MW 95] Mandel, L. und Wolf, E.: *Optical coherence and quantum optics*, Cambridge University Press, Cambridge, 1995.

[MW 98] Mahler, G. und Weberruß, V. A.: *Quantum networks – dynamics of open nanostructures*, 2 ed., Springer-Verlag, Berlin, 1998.

[NC 00] Nielsen, M. A. und Chuang, I. L.: *Quantum Computation and quantum information*, Cambridge University Press, Cambridge, 2000.

[NCS 98] Nielsen, M. A., Caves, C. M., Schumacher, B. und Barnum, H.: *Information-theoretic approack to quantum error correction and reversible measurement*; Proc. R. Soc. Lond. **A 454**, 277–304 (1998), quant-ph/9706064.

[NPN 97] Namiki, M., Pascazio, S. und Nakazato, H.: *Decoherence and quantum measurements*, World Scientific, Singapore, 1997.

[Omn 94] Omnès, R.: *The Interpretation of Quantum Mechanics*, Princeton University Press, Princeton, 1994.

[Pas 05] Passon, O.:*Why isn't every physicist a Bohmian?*; quant-ph/0412119v2.

[Pel 98] Pellizzari, T.: *Quantum computers, error-correction and networking: Quantum optical approaches*, in [LSP 98], p. 270–310, 1998.

[Per 90] Peres, A.: *Neumark's Theorem on Quantum Inseparability*, Found. Phys. **20**, 1441–1453 (1990).

[Per 93] Peres, A.: *Quantum theory – concepts and methods*, Kluwer, Dordrecht, 1993.

[Per 96] Peres, A.: *Separability criterion for density matrices*, Phys. Rev. Lett. **77**, 1413-1415 (1996); quant-ph/9604005.

[Pet 63] Petersen, A.: *The Philosophy of Nils Bohr*, Bull. Of the Atomic Scientists, **19**, 8–14 (1963).

[Pit 00] Pittenger, A. O.: *An Introduction to Quantum Computing Algorithms*, Birkhäuser, Basel, 2000.

[PR 04] Paris, M. G. A. und Řeháček, J.: *Quantum State Estimation*, Springer, Heidelberg, 2004.

[Pre 98] Preskill, J.: *Course Informations for Physics 219/Computer Science 219 Quantum Computation (Formerly Physics 229)*, California Institute of Technology, 1998, *http://www.theory.Caltech.edu/ preskill/ph229*.

[Pre 98 a] Preskill, J.: *Fault-tolerant quantum computation*, in [LSP 98], p. 213–269, 1998; quant-ph/9712048.

[Pre 99] Preskill, J.: *Battling decoherence: The fault-tolerant quantum computer*, Phys. Today **52**, 6, 24, 30 (1999).

[Pri 81] Primas, H.: *Chemistry, Quantum Mechanics and Reductionism*, Springer-Verlag, Berlin, 1981.

[PSE 96] Palma, G. M., Suominen, K. und Ekert, A. K.: *Quantum computers and dissipation*, Proc. R. Soc. Lond. **A 452**, 567–584 (1996).

[PSM 87] Prasad, S., Scully, M. O. und Martienssen, W.: *A Quantum Description of the Beam Splitter*, Optics Communications **62**, 139–145, (1987).

[PT 04] Peres, A. und Terno, D. R.: *Quantum information and relativity theory*, Rev. Mod. Phys. **76**, 93–123 (2004).

[PV 98] Plenio, M. B. und Vedral V.: *Teleportation, entanglement and thermodynamics in the quantum world*, Contemporary Physics **39**, 431–446 (1998).

[PVK 96] Plenio, M. B., Vedral, V. und Knight, P. L.: *Computers and communication in the quantum world*, Phys. World **9**, 10, 19–20 (1996).

[PZ 02] Paz, J. P. und Zurek, W. H.: *Enviroment-Induced Decoherence and the Transition from Quantum to Classical* in [Hei 02], p. 77–148, 1902.

[RBH 01] Raimond, J.M., Brune, M. und Haroche, S.: *Manipulating quantum entanglement with atoms and photons*, Rev. Mod. Phys. **73**, 565–582 (2001).

[Rem 06] Rempe, G.: *Entangled quantum systems: from wave-particle duality to single-photon sources of light*, in [Aud 06], p. 113-141, 2006.

[RKM 01] Rowe, M. A., Kielpinski, D., Meyer, V., Sackett, C. A., Itano, M., Monroe, C. und Wineland, D. J.: *Experimental violation of a Bell's inequality with efficient detection*, Nature **409**, 791-794 (2001).

[RP 00] Rieffel, E. und Polak, W.: *An Introduction to Quantum Computing for Non-Physicists*; quant-ph/98090 v2 (2002).

[RW 00] Rauch, H. und Werner, S.A.: *Neutron Interferometry: Lessons in Experimental Quantum Mechanics*, Clarendon Press, Oxford, 2000.

[Sak 85] Sakurai, Jun John: *Modern Quantum Mechanics*, Addison-Wesley, New York, 1985.

[Sch 35] Schrödinger, E.: *Die gegenwärtige Situation in der Quantenmechanik*, Naturwissenschaften **23**, S. 807–812, S. 823–828, S. 844–849 (1935). Engl. Übersetzung: Proc. American Philosophical Society **124**, 323-338 (1980).

[Sch 64] Scheibe, E.: *Die kontingenten Aussagen in der Physik. Axiomatische Untersuchungen zur Ontologie der klassischen Physik und der Quantentheorie*, Frankfurt, 1964.

[Sch 90] Schröter, J.: *Das L-Konzept physikalischer Theorien*, Praxis der Naturwissenschaften – Physik 39, Heft **2**, 20–27 (1990).

[Sch 95] Schumacher, B.: *Quantum coding*, Phys. Rev. **A51**, 2738–2747 (1995)

[Sch 96a] Schröter, J.: *Zur Meta-Theory der Physik*, Walter de Gruyter, Berlin, 1996.

[Sch 96b] Schumacher, B.: *Sending entanglement through noisy channels*, Phys. Rev. **A 54**, 2614–2628 (1996).

[Sch 04] Schlosshauer, M.: *Decoherence, the measurement problem, and interpretations of quantum mechanics*, Rev. Mod. Phys. **76**, 1267-1305 (2004).

[Sch 05] Schlosshauer, M.: *Experimental motivation and empirical consistency in minimal non-collapse quantum mechanics*; quant-ph/0506199 (2005).

[See 04] Seevinck, M. P.: *Holism, physical theories and quantum mechanics*, Studies in History and Philosophy of Science Part B, Vol. **35**, Issue 4, 693–712 (2004).

[SH 04] Steeb, W.-H. und Hardy, Y.: *Problems and solutions in quantum computing and quantum information*, World Scientific, Singapore, 2004.

[Sha 48] Shannon, C. E.: *A mathematical theory of communication*, Bell System Tech. J. **27**, 379–423, 623–656 (1948).

[Sha 49] Shannon, C. E.: *Communication theory of secrecystems*, Bell Systems Technical Journal **28**, 656–715 (1949).

[Shi 63] Shimony, A.: *Role of the observer in quantum theory*, Am. J. Phys. **31**, 755-773 (1963).

[Sho 94] Shor, P. W.: *Algorithms for quantum computation: Discrete logarithms and factoring*, in Goldwasser, S. (Ed.): *Proc. of the 35th Annual Symp. on the Foundations of Computer Science (Santa Fe, New Mexico, 1994)*, IEEE Computer Science Society Press, Los Alamitos, California, 1994, p. 124–134.

[Sho 97] : Shor, P.W.: *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, Soc. Ind. Appl. Math. J. Comp. **26**, 1484–1509 (1997); quant-ph/9508027.

[SIG 05] Scarani, V., Iblisdir, S., Gisin, N. und Acin, A.: *Quantum cloning*, Rev. Mod. Phys. **77**, 1225-1256 (2005).

[SS 04] Stolze, J. und Suter, D.: *Quantum Computing – A Short Course from Theory to Experiment*, Wiley-VCH, Weinheim, 2004.

[SS 05] Stenholm, S. und Suominen, K.-A.: *Quantum approach to informatics*, Wiley, Hoboken, N.J., 2005.

[SSS 04] Scherer, A., Soklakov, A. N. und Schack, R.: *A simple necessary decoherence condition for a set of histories*, Phys. Lett. **A 326**, 307–314 (2004).

[Sta 93] Stapp, H. P.: *A quantum theory of the mind-brain interface*, in *Mind, Matter, and Quantum Mechanics*, 145-172, Springer, Berlin, 1993.

[Sta 96] Stamatescu, I.-O.: *Stochastic Collapse Modells*, in [GJK 96], p. 249–267 (1996).

[Ste 98] Steane, A. M.: *Quantum computing*, Rep. Prog. Phys. **61**, 117–173 (1998); quant-ph/9708022.

[TBM 95] Torgerson, J. R., Branning, D., Monken, C. H. und Mandel, L.: *Experimental Demonstration of the Violation of Local Realism Without Bell Inequalities*, Phys. Lett. **A204**, 323-328 (1995).

[Ter 02] Terhal, B. M.: *Detecting Quantum Entanglement*, quant-ph/0101032 v1, (2001) und Journal of Theoretical Computer Science **287**, 313–335 (2002).

[Vai 01] Vaidman, L.: *The Many-Worlds Interpretation of Quantum Mechanics*, Stanford Encyclopedia of Physics, *http://plato.stanford.edu/entries*, 2001.

[Vai 03] Vaidman, L.: *The Meaning of the Interaction-Free Measurements*, Found. Phys. **33**, 491-510 (2003); quant-ph/0103081.

[VC 04] Vandersypen, L.M.K.: *NMR techniques for quantum control and computation*, Rev. Mod. Phys. **76**, 1037 (2004).

[VdW 00] van der Wal, C. H., ter Haar, A. C. J., Wilhelm, F. K., Schouten, R. N., Harmans, C. J. P. M., Orlando, T. P., Lloyd, S. und Mooij, J. E.: *Quantum superposition of macroscopic persistent-current states*, Science **290**, 5492, p. 773–777 (2000).

[Ved 02] Vedral, V. M: *The role of relative entropie in quantum information theory*, Rev. Mod. Phys. **74**, 197–234 (2002)

[Ver 26] Vernam, G. S.: *Cipher printing telegraph systems for secret wire and radio telegraphic communication*, J. Am. Just. Electr. Eng. **45**, 109–115 (1926).

[vLo 02] Van Loock, P.: *Quantum Communication with Continuous Variables*, Fortschr. Phys. **50**, 1177–1372 (2002).

[vNe 68] von Neumann, J.: *Mathematische Grundlagen der Quantenmechanik*, Springer-Verlag, Berlin, 1968.

[VP 98] Vedral, V. und Plenio, M. B.: *Basics of quantum computation* ; quant-ph/9802065.

[VRP 97] Vedral, V., Rippin, M. A. und Plenio, M. B., *Quantum correlations, local interactions and error correction*, J. of Mod. Optics **44**, 2185–2205 (1997).

[VV 00] Vojta, G. und Vojta, M.: *Teubner Taschenbuch der statistischen Physik*, Teubner, Stuttgart, 2000.

[WC 98] Williams, C. P. und Clearwater, S. H.: *Explorations in quantum computing*, Springer-Verlag, New York, 1998.

[Weh 78] Wehrl, A.: *General properties of entropy*, Rev. Mod. Phys. **50**, 221–260 (1978).

[Wei 06] Weinfurter, H.: *Quantum information*, in [Aud 06], p. 143-168, 2006.

[Wer 89] Werner, R. F.: *Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model*, Phys. Review **A 40**, 4277–4281 (1989).

[Wer 01] Werner, R. F.: *Quantum Information Theory – an Invitation*, in [ABH 01], p. 14–57.

[Wer 06] Werner, R.F.: *Quantum computers – the new generation of supercomputers*, in [Aud 06], p. 169-201, 2006.

[WH 02] Walgate, J. und Hardy, L.: *Nonlocality, asymmetry and distinguisihing bipartite states*, Phys. Rev. Lett. **89**, 147901 (2002); quant-ph/0202034.

[Whe 78] Wheeler, J. A.: *The 'Past' and the 'Delayed Choice' Double Slit Experiment*, in [Mar 78], p. 9–48, 1978.

[Whe 90] Wheeler, J. A.: *Information, Physics, Quantum: The Search for Links* in Zurek, W. H. (Ed.): *Complexity, Entropy and the Physics of Information*, S. 3-28, Addison-Wesley, New York, 1990.

[Wie 83] Wiesner, S.: *Conjugate coding*, Sigact news, 15 (1), 78–88 (1982).

[WMN 98] White, A. G., Mitchell, J. R., Nairz, O. und Kwiat, P. G.: *"Interaction-free" imaging*, Phys. Rev. **A58**, 605-613 (1998); quant-ph/9803060.

[Woo 98] Wootters, W. K.: *Entanglement of formation of an arbitrary state of two qubits*, Phys. Rev. Lett. **80**, 2245-2248 (1998); quant-ph/970929.

[Woo 01] Wootters, W. K.: *Entanglement of formation and concurrence*, Quant. Inf. Comp. **1**, 27-44 (2001).

[WSH 00] Walgate, J., Short, A. J., Hardy, L. und Vedral, V.: *Local distinguishability of multipartite orthogonal quantum states*, Phys. Rev. Lett. **85**, 4972-4975 (2000); quant-ph/0007098.

[WW 00] Werner, R. F. und Wolf, M. M.: *Bell's inequalities for states with positive partial transpose*, Phys. Rev. **A61**, 062102 (2000); quant-ph/9910063.

[WW 01] Werner, R. F. und Wolf, M. M.: *Bell Inequalities and Entanglement*, Quantum Information and Computation **1**, 1–25 (2001).

[WZ 82] Wootters, W.K. und Zurek, W.H.: *A single quantum cannot be cloned*, Nature **299**, 802-803 (1982).

[Zbi 98] Zbinden, H.: *Experimental Quantum Cryptography*, in [LSP 98], p. 120–142, (1998).

[Zeh 70] Zeh, H. D.: *On the interpretation of measurement in quantum theory*, Found. Phys. **1**, 69-76 (1970).

[Zeh 96] Zeh, H. D.: *The Program of Decoherence: Ideas and Concepts*, in [GJK 96], p. 5–34 (1996).

[Zeh 00] Zeh, H. D.: *The Meaning of Decoherence*, in [BGJ 00] p. 19–42, (2000).

[Zei 81] Zeilinger, A.: *General properties of lossless beam splitters in interferometry*, Am. J. Phys. **49**, 882-883 (1981).

[Zur 81] Zurek, W. H.: *Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse?*, Phys. Rev. **D24**, 1516-1525 (1981).

[Zur 82] Zurek, W. H.: *Environment-induced superselection rules*, Phys. Rev. **D 26**, 1862–1880 (1982).

[Zur 91] Zurek, W. H.: *Decoherence and the transition from quantum to classical*, Phys. Today **44**, 10, 36–44 (1991).

[Zur 02] Zurek, W. H.: *Decoherence and the Transition from Quantum to Classical-Revisited*, Los Alamos Science, Nr. **27**, 2–25 (2002).

[Zur 03] Zurek, W. H.: *Decoherence, einselection, and the quantum origins of the classical*, Rev. Mod. Phys. **75**, 715–775 (2003).

[ZZH 93] Zukowski, M., Zeilinger, A., Horne, M. A. und Eckert, A. K.:*„Event-ready detectors" Bell experiment via entanglement swapping*, Phys. Rev. Lett. **71**, 4287-4290 (1993).

© 2007 WILEY-VCH Verlag GmbH & Co.

Subject Index

50 : 50 beam splitter **65**
$\sqrt{\text{NOT}}$gate 60
2-qubit gate 242

addition theorem
    for probabilities **18**
additivity **169**
additivity axiom **16**
Alice **116**
alternative theories 42, **43**, 188, 191
amplitude amplification **225**
amplitude damping channel 283
ancilla **248**
ancilla system 261, 310
angular-momentum operators **51**
antilinear **2**
Araki-Lieb inequality **171**
arithmetic, modular **221**

B92 protocol **200**
basic domain **41**
basis **117**
    magic **214**
    orthonormal **3**
Bayes' assumption **19**, 269
Bayes' theorem 18, 76, 269
BB84 protocol **201**
BBM92 protocol **203**
beam splitter, lossless **64**
Bell basis **174**
Bell diagonal state **217**
Bell inequality **189**, 190, 203
Bell measurement **180**, 207
Bell state **118**, 158, 179, 184, 204, 205, 207
Bell's theorem **202**
Bell-CHSH observable 192
bi-orthogonal decomposition 260
bi-orthogonal expansion **149**
bilateral 209
binary coding 92
binary digits **93**
binary string 93
bipartite system **115**
bit by bit measurements **220**
bit flip **52**
bit flip error **239**
bit sequence **93**
bits **92**, 93
black box **225**
blend **26**, 31, 75, 302
Bloch
    point **53**
    sphere **53**, 281, 291
    vector **53**, 85
block coding 95
Bob **116**
Bohm theory **196**
Boolean function *see* function, Boolean
bra space **2**
bra vector **2**, 117
branch **304**

cascade **157**
cavity QED **241**
CC *see* classical communication
CCNOT gate **137**
character string **90**
choice, delayed **166**
CHSH inequality **190**, 203
circuit **66**
    quantum **66**
classical communication **144**
classical mixture 25
classical physics **293**

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch
Copyright © 2007 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-40684-5

classical state 295
classically correlated **173**
classically correlated quantum state **144**
CNOT gate **136**, 239
CNOT transformation **179**
code, quantum error-correcting **239**
coding, binary **92**
collapse and revival **301**
collapse theory **306**
collapse, of the wavefunction 35
communication, classical **144**
complement, orthogonal **3**
complete positivity 274
complete quantum operation 310
complete system of commuting observables 10
completeness relation **5**
component **3**
  mathematical 41
composite quantum system **120**
compound system *see* composite system
computational basis **49**
computational complexity 244
concurrence **212**
conditional entropy **99**
conditional probability 17
conjugate linear **308**
consciousness **305**
conservation of the norm **11**
consistent histories **306**
continuous-variable entanglement **46**
control pair **210**
control qubit **136**
controlled NOT gate *see* CNOT gate
controlled-controlled NOT 137
convex sum 79
Copenhagen interpretation **44**
copy **201**
correlation **99**, 100
  degree of 170
correlation at a distance 187
correlation coefficient **185**
correspondence rules **41**
  hypothetical **42**
correspondence, dual **2**
cryptogram **199**

data compression **96**
de Broglie-Bohm theory *see* Bohm theory
decoherence 80, 239, 244, 290–292
  environment-induced **293**, 295, 297
decomposition
  bi-orthogonal **9**, **149**, 260
  dyadic **5**
    of the identity operator **5**
  left-polar **259**
  orthogonal **7**
  polar **149**
  right-polar **260**
  spectral **7**
  tri-orthogonal 300
decomposition of the identity *see* dyadic decomposition of the identity operator
decomposition operator 249, **275**
degeneracy **36**
degenerate 7
degree of degeneracy **32**
degree of mixtureness 79, 85
delayed choice 46, **165**, 166
delayed entanglement 209
density matrix **73**
density operator **73**, 75, 79
  maximally mixed **79**
  reduced 82, 122, **122**
depolarisation **282**
determination of states **85**
Deutsch gate **138**
Deutsch problem **225**
Deutsch-Jozsa problem **225**
developed domain of reality **42**
deviation, mean square **19**
dimension, of Hilbert space **35**
Dirac notation **2**
dispersion **19**
distillation protocol 209
divisor, greatest common **231**
domain of reality 41
dual correspondence 2
dyad **5**
dyadic decomposition 5
dynamics
  unitary **33**, 55, 56, 80, 108, 131, 247, 256

eavesdropping **201**
effect operator **260**, *see* POVM element
eigenbasis **7**, 105

eigenvalue **3**
eigenvector **3**
Einstein
  locality 135, **188**
  reality 135, **188**
either-or states **25**
elements of physical reality **188**
emergent property 294
ensemble **75**
ensemble decomposition **83**, 110
entangled **117**, 224, 248
entangled quantum state **145**
entanglement **116**, 143, 152, 171, 253
entanglement distillation **209**
entanglement of formation **214**
entanglement swapping **207**
entanglement witness **193**
entropy
  classical **91**
  conditional 99
  of entanglement **153**
  operational interpretation of **96**
  relative **97**
  Shannon's 89
  theoretical relative **109**
  von Neumann 89
environment-induced decoherence 295
EPR **145**
EPR correlations **145**, 185
error **96**
error correction protocol **201**
estimating a state **269**
Euclidian algorithm **243**
Eve **201**
events
  exclusive **16**
  independent **17**
  random **15**
Everett interpretation **303**
exchange coupling 141
expectation value **6**
experimentum crucis **191**
extended operator **118**

factor space **117**
fidelity **96**, **210**
forerunner theories **41**
form, non-normalised **74**
free system 31
frequency, relative **16**
fringe contrast **68**
function, Boolean **221**

general assumption **23**
general measurement 256
general non-selective measurement 257
general selective measurement 279
generalised selective measurement 251, **257**
GHZ state **194**
Gleason's theorem 82, **86**
greatest common divisor *see* divisor, greatest common
Grover algorithm **225**

Hadamard
  beam splitters **66**
  gate **60**, 226
  transformation 179
Hamiltonian **34**
Hermitian operators 9
Hilbert space, finite-dimensional **1**

idempotent **12**
identical particles *see* particles, identical
identity operator **3**
ignorance interpretation **84**, 302
implementation of the Hadamard gate
  optical **65**
improper mixture **123**
in-out approach **132**, 273
incoherent superposition 80
incomplete **188**
index state **307**
indistinguishability **116**
individual identity **116**
inequality
  Gibbs' **97**
  Schwarz **2**
  triangle **171**
inequality, Bell 189
inference 18
information content **96**
information transmission
  assisted by entanglement 205
information, mutual **98**, 170, 174
informationally complete **268**
informationally complete measurement **265**
inner product 2

integer **221**
interference **67**
interference pattern **24**, 67, 81
interpretation **41**
    London-Bauer **305**
intervention **256**, 273
intrinsic properties 35
inverse operator 3
ion trap **241**

joint entropy **169**
joint probability 15, **98**

ket space **2**
ket vector **2**
key **199**, 200
kick back **224**
Klein's inequality **109**
Kraus operator **249**, 275, 290

law of large numbers 94
left-polar decomposition 259
Lindblad master equation **278**
Lindblad operator **278**
Lindblad superoperator 278
Lindbladian **278**
linear operator 3
Liouville
    space **13**
Liouville operator **15**, 74, 131
Liouvillian 15
LO *see* local operation
local **115**, 187, 188
local measurement **122**, **125**
local observable **122**
local operation 115, **144**
LOCC **144**
loophole 196
lossless beamsplitter 64

Mach-Zehnder interferometer **66**, 161
magic basis 214
manner of speaking 30
many-worlds interpretation 45, **303**
mapping principles **41**
marker
    observable **164**
    state **163**, 307
    system **163**
marking **290**
Markovian approximation **279**
master equation **277**
matrix elements **5**
maximally entangled state **153**
mean square deviation 19
mean value 6, 10, **19**
measure of entanglement **152**
measured values, correlated **130**
measurement **105**
    general 256
        selective 256
    general selective **279**
    generalised 178, 262
        selective **251**, **257**
    informationally complete 265
    local 121, **122**, 125, 128, 175
    minimal **260**, 261
    most general selective **311**
    non-local 175, 239, 240
    non-selective **77**, 110, 303
    projective *see* projective measurements
    selective 26, **76**, 303
    sharp **255**
    unsharp **255**
measurement dynamics 55, **247**, 256
measurement operator **250**, **257**
measurement outcome 257, 285
measurement-based model 244
measuring device **285**
memory state **304**
message **90**
microtraps 141
minimal interpretation 42, **43**, 84
minimal measurement 260
mixture **82**, 123
    separable **172**
    statistical **25, 26**, 31, 75
model
    measurement-based **244**
modular arithmetic, *see* modular arithmetic
most general selective measurements 311
mutual information 98

natural philosophy **40**
negative-result measurement 68
neither-nor state **26**
no-cloning theorem **161**
non-local measurement 175

non-local observable 175
non-selective measurement 27
norm **2**
normal operator 6
NOT gate 60
null measurement **68**
number
  natural **221**

observable **10**, 32
  Bell-CHSH **192**
  classical **44**
  collective **121**
  non-local **121**, 175
ONB **3**
ontology **40**
open quantum system 131
open system *see* open quantum system
operation element **249**, 275
operation, local **115**, **144**
operational 93, **116**
operator
  σ *see* Pauli operator
  diagonalisable **6**
  extended **118**
  function **9**
  Hermitian **13**
  inverse **3**
  linear **3**
  local **122**
  normal **6**
  unitary **11**
operator basis **50**
  orthonormal 14
operator-sum decomposition **249**, 275
operator-sum representation **249**
oracle **225**
orthogonal **2**, 14
orthogonal complement **3**
outer product *see* dyadic product
outer product representation *see* dyadic decomposition
output text **93**

parallelogram equation **3**
parity bit **179**
partial trace **119**
partial transposition **155**
particles
  identical **139**
Pauli
  matrices **52**
  operator 49, **51**, 175
Peres-Horodecki criterion *see* PPT criterion
phase bit **179**
phase damping channel **291**
phase flip **52**
phase flip channel **292**
phase flip error **240**
phase gate **60**
phase shifter **60**, 81
philosophy of science **40**
photon **200**, 204
photon polarisation **62**
physical properties 34
pointer 298
pointer observable 302
pointer state **302**
polar expansion **149**
positive integer 221
positive operator 11
positive operator valued measure, *see* POVM
positive partial transpose criterion *see* PPT criterion
positivity, complete **274**
postulates **32**, 82, 121, 302
POVM **260**
  element **260**, 264
  measurement **264**, 266, 267, 280
PPT criterion 156
pre-measurement **298**
preparation
  of a state **105**
preparation apparatus **285**
preparation entropy **111**
preparation procedure **25**, 30, 36
prime factorisation **231**
private amplification protocol **201**
probabilistic cloning **167**
probability 16, 33, 122
  classical **75**
  conditional 76, 98
product
  dyadic **5**
  Hilbert space **116**
  inner **2**
  operator **118**
  vector **117**

projection operator 12
  local **122**
projection valued measure **264**
projective measurements **32**, 38, 55
projector **7**
proper mixture **123**
property
  emergent **294**
  intrinsic **139**
protocol **200**
pure state 31
purification **152**, **209**
PVM **264**

quantisation, second 141
quantum algorithm **225**
quantum channel **103**, 281
quantum circuit 135, **179**, 220
quantum code, error correcting **244**
quantum coding, dense **204**
quantum computation **220**
quantum computers 136, **220**
quantum correlations **145**
quantum data compression **107**
quantum dense coding 204
quantum dots 141, **241**
quantum effects **29**
quantum entropy **105**
quantum erasure **164**
quantum error-correcting code 239
quantum Fourier transform 236
quantum gates **49**, 59, 135, 220
  universal 66, **138**, **241**
quantum information **106**, 112
quantum information theory 89
quantum jumps **282**
quantum measurement process **42**
quantum network **220**
quantum objects **45**
quantum operation **274**, 275, 290, 311
  complete 275, **310**
  incomplete 275
quantum parallelism **224**
quantum range **29**
quantum register **219**
quantum signal ensemble **103**
quantum signal source **103**
quantum state 26, **30**, 82
  classically-correlated **144**
  entangled **145**
quantum state estimation 270
quantum system **29**
  composite **120**
  isolated **31**
  open **131**
quantum teleportation **205**
quantum theory **29**
quantum wires **135**, 220
quantum Zeno effect **39**, 71
quantum-to-classical transition **293**
qubit **49**, 105, 106
  system **49**, 62
query **225**

Rabi
  frequency **58**
  oscillations **58**
random variable **19**, **90**
randomised algorithm **233**
re-preparation **27**
realism, local **187**, 188
reality **29**, **40**, 44, **187**
  physical **42**
recoherence 295, **301**
reduced density operator **122**
reflection factor **65**
remainder **221**
right-polar decomposition 260

scalar product **2**, 14, 117
scenario of quantum mechanics 23, 285
Schmidt
  bases **151**
  coefficients **150**
  decomposition **149**
  number **151**
Schrödinger
  representation **34**, 131, 274
Schrödinger's cat **296**
Schumacher's quantum noiseless coding theorem **107**
second quantisation 141
secret **202**
selective measurement 76
self-adjoint **9**
separability problem **147**
separable **175**
separable state **145**

sequence **90**
  typical 91
Shannon's entropy 89, **91**, 96, 105, 110
Shannon's noiseless coding theorem *see* Shannon's theorem
Shannon's theorem **96**
signal ensemble **90**
signal source **90**
  stochastic memoryless **90**
signal state 103
simulation **123**
single-qubit gate 242
singlet state **118**
source text **199**
spectral theorem **10**
spin **51**, **62**
spin-$\frac{1}{2}$ particles 62, **203**
spooky action into the past 209
standard basis **49**
standard deviation **19**
standard interpretation **29**, 42, 44, 45, 303
state 25, **74**, 123, 224
  classical **25**, 295
  correlated **144**
  entangled **115**
  maximally entangled **153**
  pure **26**, 31, 73, 79, 80, 155
  relative **124**, 308
  separable **145**
  Werner 284
state evolution
  deterministic 257
  non-deterministic 257
state reduction 35
state vector **32**, 36
statistical mixture 25, 75, 302
Stern-Gerlach experiment **251**
  ideal **254**
  non-optimal **253**
subadditivity **101**, 170
subjectivistic interpretation **45**
subspaces
  decoherence-free **244**
subsystem **115**, 120
subsystem operator **118**
sum, convex **79**
super selection rules **44**
superoperators **14**, 279
superposition 117
supremum norm **6**
SWAP gate **137**
system
  closed 31
  composite 82, **115**
  open **31**, **247**

target pair **210**
target qubit **136**
tensor product **116**, **120**
theorem
  of the operator-sum decomposition 275
theory
  classical **188**
  physical 41
  stochastic, local-realistic **188**
time **35**
time development operator 33
Toffoli gate **137**
total probability 18
trace **6**, 119
transfer of entanglement **207**
transformation apparatus **28**, 285
transmission factor **65**
transposition **155**
trapped atoms **241**
tri-orthogonal decomposition 300
triangle inequality **2**, **171**
two-slit experiment **23**
typical subspace **107**

U gate, controlled **137**
uncertainty **19**, **91**
  remaining **100**
unilateral 209
unitary 11
unitary equivalence **11**
universal quantum gates 138, 241
unsharp measurement 255

values
  of a measurement **32**
variables, classical **35**
variables, hidden **188**
variance **19**
von Neumann
  entropy **89**, 111
  equation **74**, 131

Werner state 284
which-way marker **163**

XOR gate *see* CNOT gate

Zeno
    effect **39**
    time **39**